University Science Books
55D Gate Five Road
Sausalito, CA 94965

Fax: (415) 332-5393
www.uscibooks.com

Production manager: *Susanna Tadlock*
Manuscript editor: *Jeannette Stiefel*
Designer: *Robert Ishi*
Compositor: *Eigentype*
Printer & Binder: *Maple-Vail Book Manufacturing Group*

This book is printed on acid-free paper.

Library of Congress Cataloging-in-Publication Data

Bloomfield, Victor A.
Nucleic acids : structures, properties, and functions / by Victor A. Bloomfield, Donald M. Crothers, Ignacio Tinoco, Jr. ; with contributions from John E. Hearst ... [et al.].
p. cm.
Includes bibliographical references and index.
ISBN 0-935702-49-0 (hardcover)
1. Nucleic Acids. I. Crothers, Donald M. II. Tinoco, Ignacio. III. Title.
QP620.B64 1999
572.8—dc21 98-45268
CIP

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

NUCLEIC ACIDS

STRUCTURES, PROPERTIES, AND FUNCTIONS

Victor A. Bloomfield
UNIVERSITY OF MINNESOTA

Donald M. Crothers
YALE UNIVERSITY

Ignacio Tinoco, Jr.

WITH CONTRIBUTIONS FROM

John E. Hearst
UNIVERSITY OF CALIFORNIA, BERKELEY

David E. Wemmer
UNIVERSITY OF CALIFORNIA, BERKELEY

Peter A. Kollman
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Douglas H. Turner
UNIVERSITY OF ROCHESTER

University Science Books
Sausalito, California

Ravishanker, G., Swaminathan, S., Beveridge, D. L., Lavery, R., and Sklenar, H. (1989). Conformational and Helicoidal Analysis of 30 PS of Molecular Dynamics on the d(CGCGAATTCGCG) Double Helix: Curves, Dials and Windows. *J. Biomol. Struct. Dyn.* **6**, 669–699.
Saenger, W., (1984). *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Sasisekharan, V., (1973). *Conformation of Biological Molecules and Polymers*, Bergmann, E. and Pullman, B., Eds., Jerusalem, Israel.
Schaefer III, H. F. (1986). Methylene, A Paradigm for Computational Quantum Chemistry. *Science* **231**, 1100–1107.
Seeman, N., Rosenberg, J., and Rich, A. (1976). Sequence-Specific Recognition of Double Helical Nucleic Acids by Proteins. *Proc. Nat. Acad. Sci.* **73**, 804–808.
Sharp, K. A. and Honig, B. (1990). Electrostatic Interactions in Macromolecules: Theory and Experiment *Annu. Rev. Biophys. Biophys. Chem* **19**, 301–332 (1990).
Simmerling, C., Miller, J., and Kollman, P.A. Combined Locally Enhanced Sampling and Particle Mesh Ewald as a Strategy to Locate the Experimental Structure of a Non-Helical Nucleic Acid. *J. Am. Chem. Soc.* (in press).
Sinden, R. and Hagerman, P. (1984). Interstrand Psoralen Cross-Links Do Not Introduce Appreciable Bends in DNA, *Biochemistry* **23**, 6299–6303.
Singh, U. C. and Kollman, P. (1984). An Approach to Computing Electrostatic Charges for Molecules. *J. Comp. Chem.* **5**, 129–145.
Singh, U. C. and Kollman, P. (1985). A Combined *Ab Initio* Quantum Mechanical and Molecular Mechanical Method for Carrying out Simulations on Complex Molecular Systems: Applications to the $CH_3Cl^+Cl^-$ Exchange Reaction and Gas Phase Protonation of Polyethers. *J. Comp. Chem.* **7**, 718–730.
Singh, U. C., Pattabiraman, N., Langridge, R., and Kollman, P. A. (1986). Molecular Mechanical Studies of $d(CGTACG)_2$: Complex of Triostin A with the Middle A·T Base Pairs In Either Hoogsteen or Watson–Crick Pairing. *Proc. Natl. Acad. Sci. USA* **83**, 6402–6406.
Singh, U. C., Weiner, P., Caldwell, J., and Kollman, P. (1986), *AMBER 3.0* University of California, San Francisco, CA.
Srinivasan, J., T. E. Cheatham, P. Cieplak, P. A. Kollman and D. A. Case (1998). Continuum Solvent Studies of the Stability of DNA, RNA and Phosphoramidate-DNA Helices, *J. Amer. Chem. Soc.*, **120**, 9401–9409
Still, W. C., Tempczyk, A., Hawley, R. C., and Henrickson, T. (1990). Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *J. Amer. Chem. Soc.* **112**, 6127–6120.
Subramanian, P. S., Swaminathan, S., and Beveridge, D. L. (1990). Theoretical account of the 'spine of hydration' in the minor groove of duplex d(CGCGAATTCGCG). *J. Biomol. Struct. Dyn.* **7**, 1161–1165.
Swaminathan, S., Ravishanker, G., and Beveridge, D. (1991). Molecular dynamics of B-DNA including water and counterions: A 140-ps trajectory for d(CGCGAATTCGCG) based on the GROMOS force field. *J. Am. Chem. Soc.* **113**, 5027–5040.
Szabo, A. and Ostlund, N. (1982). *Modern Quantum Chemistry*, Macmillan, New York, (1982).
Tilton R. F., Jr., Weiner, P. K., and Kollman, P. A. (1983). An Analysis of the Sequence Dependence of the Structure and Energy of A- and B-DNA Models Using Molecular Mechanics. *Biopolymers* **22**, 969–1002.
Van Gunsteren, W. F. and Berendsen, H. J. C. (1990). Molecular Dynamics Computer Simulation. Method, Application and Perspectives in Chemistry. *Angew. Chem* **102**, 1020–1055.
Veillard, A., Ed. *Quantum Chemistry: The Challenge of Transition Metals and Coordination Chemistry*, Vol. 176, NATO ASI Series C Dordecht, The Netherlands.
Warshel, A. and Levitt, M. (1976). Theoretical Studies of Enzymatic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **103**, 227–249.
Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., and Weiner, P. (1984). A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc.* **106**, 765–784.
Weiner, P. K., Landridge, R., Blaney, J. M., Schaefer, R., and Kollman, P. A. (1982). Electrostatic Potential Molecular Surfaces. *Proc. Natl. Acad. Sci. USA* **79**, 3754–3758.
Woodward, R. B. and Hoffman, R. (1970). *The Conservation of Orbital Symmetry*, Verlag Chemie, Weinheim, Bergstrasse.
York, D. M., Wlodawer, A., Pedersen, L. G., and Darden, T. A. (1994). Atomic Level Accuracy in Simulations of Large Protein Crystals. *Proc. Natl. Acad. Sci.* **91**, 8715–8718.

CHAPTER 8

Conformational Changes

by Douglas H. Turner

To be functional, nucleic acids must adopt particular three dimensional (3D) conformations. For example, Figure 8-1(*a*) shows the sequence of a self-splicing intron from *Tetrahymena thermophila* (Cech and Bass, 1986). At high temperatures, this molecule has the conformation of a random coil, and is not active. At temperatures near 37°C, however, the interactions discussed in Chapter 7 force the molecule to fold on itself to give the base pairing, or secondary structure, shown in Figure 8-1(*b*) (Cech

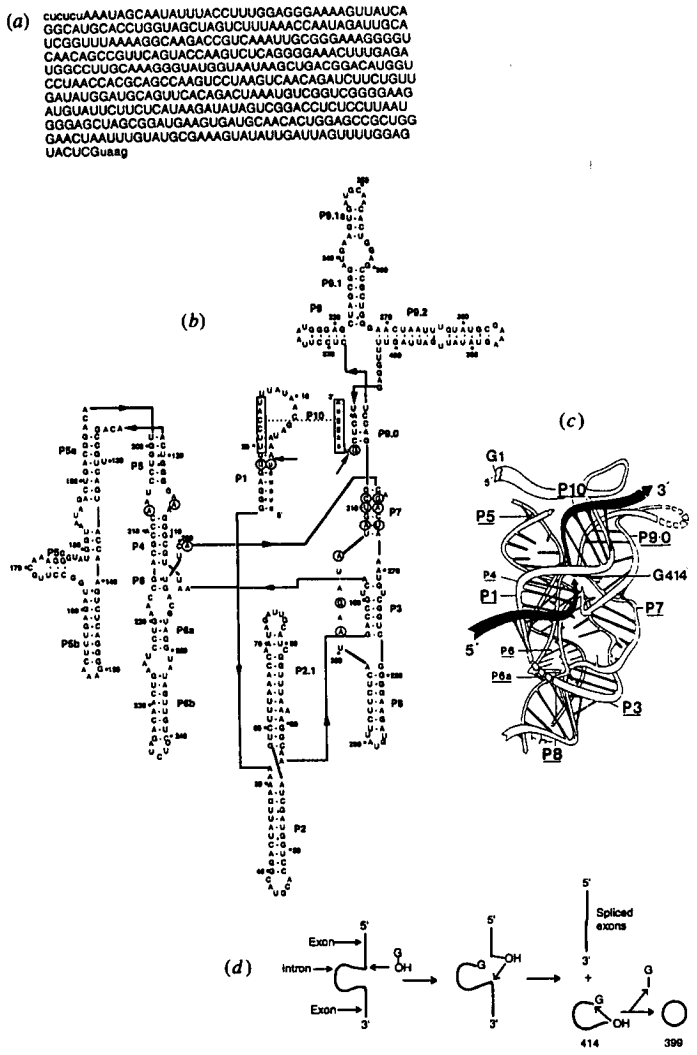

Figure 8-1 (*facing page*).
(*a*) The sequence of the self-splicing intron from the rRNA precursor of *T. thermophila* is shown in capital letters. Small letters give the adjacent sequences of the attached exons. (*b*) Secondary structure formed by the sequence in (*a*). [Reprinted with permission from Cech, R. R., Damberger, S. H., and Gutell, R. R. (1994). *Nature Struct. Biol.* **1**, 273–280.] Nucleotides that are circled are conserved in other group I introns. (*c*) Proposed model for 3D structure of active site for self-splicing intron [Reprinted with permission from Michel and Westhof, 1990]. (*d*) Sequence of reactions mediated by structure in (*b*). [From Cech and Bass, 1986.]

et al., 1994). These same interactions also determine the 3D folding of the molecule, which has been suggested to look like Figure 8-1(*c*) at the active site (Michel and Westhof, 1990). When the molecule is in the necessary conformation, it can catalyze the reactions shown in Figure 8-1(*d*) (Cech and Bass, 1986; Cech, 1990). Since transcription of RNA occurs by stepwise addition of nucleotides, conformational changes are required for it to assume its active state.

Self-splicing introns are examples of nucleic acids with catalytic functions. The function of much nucleic acid, however, is to carry information, such as the sequence of a protein. In these cases, conformation is still important for interactions with other molecules that regulate expression and transmission of the information (Oxender et al., 1979; Shen et al., 1988; Tang and Draper, 1989).

As described in Chapter 3, it is now straightforward to determine the sequence of a nucleic acid. Sequences containing more than 1 billion nucleotides are known, and determination of the 3 billion nucleotides of the human genome is in sight. Determination of the secondary and 3D structures and of the dynamics of these structures is more difficult, however (see Chapters 4 and 5). This finding is particularly true for RNA. Thus, transfer RNA is the only complete natural RNA whose 3D structure is known. While DNA overwhelmingly occurs as a fully paired, right-handed double helix, X-ray and solution studies reveal considerable sequence-dependent differences in conformation (Saenger, 1984; Dickerson et al., 1982; Wang et al., 1982; Arnott et al., 1982; Calladine, 1982; Dickerson, 1983). Presumably, these differences are important for interactions of DNA with other molecules. Thus the sequence dependence of conformation and its functional effects are important, but relatively unexplored.

One of the goals of biophysical chemistry is to use the available information on the structure and dynamics of nucleic acids to deduce principles that allow accurate prediction of these properties from sequence alone. In this chapter, experimental results on conformational changes are described, and the disciplines of thermodynamics, statistical mechanics, and kinetics are used to derive general principles. The results can be used to test predictions made from our knowledge of intermolecular and intramolecular forces as discussed in Chapter 7. Knowledge of these forces will ultimately permit accurate predictions of properties from sequence.

1. SINGLE-STRAND STACKING

Much biological activity of nucleic acids depends on regions of single-stranded, unpaired nucleotides. This configuration maximizes the number of active groups available

for recognition or reaction. For example, in tRNA (see Figs. 4-11–4-14), recognition of the codon depends on pairing with the single-stranded anticodon. Transfer RNA is charged with an amino acid at one of the hydroxyl groups of the 3′-terminal A in the single-stranded sequence NCCA. As shown in Figure 8-1, several single stranded nucleotides of the self-splicing intron are conserved. Presumably they are important for catalysis (Kim and Cech, 1987; Michel and Westhof, 1990).

One of the simplest conformational changes in nucleic acids involves the transition of an unpaired single strand from a random coil in which the bases are not stacked to an ordered, helical structure in which the bases are stacked. This transition is illustrated in Figure 8-2. The thermodynamics and dynamics of this transition have been studied for a number of cases, and provide insight into the interactions governing the properties of single strands. Most of the methods used in these studies are also applicable to other conformational changes.

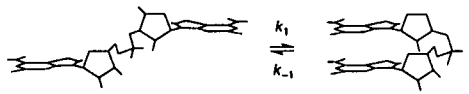

Figure 8-1
Transition from unstacked to stacked conformation in single-stranded dinucleoside monophosphate.

1.1 Thermodynamics of Single Strand Stacking

The main methods for determining thermodynamic parameters for nucleic acid conformational changes are calorimetry and the temperature dependence of spectroscopic properties. Both methods are easy to understand in principle. In practice, neither is easy to use with transitions for single-strand stacking because the transitions occur over a wide temperature range. For example, reported enthalpy changes, $\Delta H°$, for single strand stacking in poly (A) range from -3 to -13 kcal mol^{-1}.

1.1.1 Calorimetry

Calorimetric methods measure the heat released or absorbed by a chemical reaction. The most popular calorimetric method applied to conformational changes of nucleic acids is differential scanning calorimetry (DSC) (Sturtevant, 1987; Breslauer et al., 1992). In this method, electrical energy is used to slowly increase the temperatures of two matched cells. One cell contains the sample of interest and the other contains a blank (e.g., buffer). If the temperature change induces a chemical reaction in the sample but not the blank, then the amount of electrical energy required to raise the temperature of the sample cell will be increased by the amount of heat absorbed by the reaction. Thus a differential scanning calorimeter measures the difference in heat

required to raise the temperatures of the two cells. The data is reported as excess heat capacity, ϕC_p, versus temperature as shown in Figure 8-3.

One disadvantage of DSC is that any reaction occurring in the sample cell can affect the excess heat capacity. For example, if hydration of the polymer chain is temperature dependent, then ϕC_p will be affected over the temperature range where hydration is changing. This disadvantage is not serious for studies of conformational changes that occur over small temperature ranges. In these cases, baselines before and after the transition of interest can be extrapolated to subtract out from ϕC_p any effects from other reactions. This subtraction leaves the change in heat capacity, ΔC_p°, associated with the transition of interest. The area under the ΔC_p° curve for the temperature

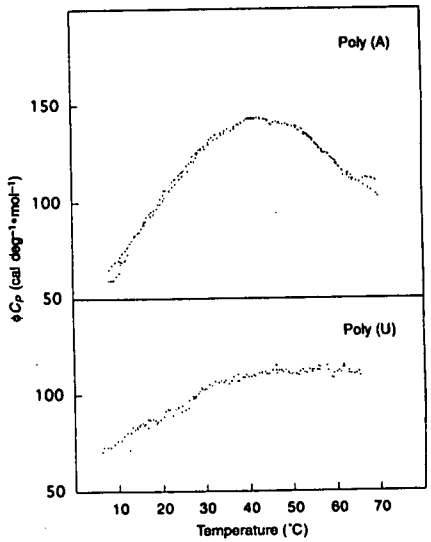

Figure 8-2
Excess heat capacity, ϕC_p, versus temperature for poly(A) and poly(U) as measured by differential scanning calorimetry. [Calorimetric Determination of the Heat capacity Changes Associated with the Conformational Transitions of Polyriboadenylic Acid and Polyribouridylic Acid, Suurkuusk, J., Alvarez, J., Freire, E., and Biltonan, R., *Biopolymers*, **16**, 2641–2652. Copyright ©1997. Reprinted by permission of John Wiley & Sons, Inc.]

interval, T1 to T2, in which the reaction is occurring is the $\Delta H°$ for the transition (see Figs. 8-3 and 8-4):

$$\Delta H° = \int_{T_1}^{T_2} \Delta C_p^\circ dT \tag{8-1}$$

Unfortunately, the stacking reaction for a single stranded polynucleotide is incomplete in the normally accessible temperature range of 0–100°C. This limitation makes it difficult to determine the baseline relevant for measuring area under the excess heat capacity curve.

1.1.2 *Temperature Dependence of Spectroscopic Properties*

The $\Delta H°$ for a reaction is related to the temperature dependence of the equilibrium constant by the van't Hoff equation:

$$\frac{\partial \ln K}{\partial T} = \frac{\Delta H°}{RT^2} \tag{8-2}$$

Thus any method that provides an equilibrium constant as a function of temperature can be used to determine $\Delta H°$. Spectroscopic methods are most common. For example, the equilibrium constant for the reaction shown in Figure 8-2 is

$$U \rightleftharpoons S \qquad K = \frac{[S]}{[U]} = \frac{\alpha}{1-\alpha} \tag{8-3}$$

where [S] and [U] are concentrations of stacked and unstacked species, respectively, and $\alpha = [S]/([S] + [U])$, the fraction of strands in the stacked state. Therefore any property that allows determination of [S] and [U] can be used to determine K. For example, the extinction coefficients for stacked and unstacked conformations are usually different (see Chapter 6, Section 1.1). Thus the absorbance, A, of a sample containing dinucleoside monophosphate will be

$$A = \varepsilon_s[S]\ell + \varepsilon_u[U]\ell = C_T\ell[\alpha\varepsilon_s + (1-\alpha)\varepsilon_u] \tag{8-4}$$

Here ε_s and ε_u are extinction coefficients for the stacked and unstacked species, respectively, and ℓ is the pathlength of the cell. If ε_s and ε_u are known, then the relevant concentrations can be determined from the absorbance and knowledge of the total concentration, $C_T = [S] + [U]$. Unfortunately, it is often difficult to determine ε_s and ε_u, particularly when the transition occurs over a large temperature range so that the sample is never completely one species.

When ε_s and ε_u cannot be directly measured, plots of absorbance versus temperature must be fit to a model for the transition of interest. For the transition shown in Figure 8-2, the two-state model is the simplest and most common model. The dinucleoside monophosphate is assumed to be either stacked or unstacked, and the $\Delta H°$ and $\Delta S°$ for the reaction are assumed to be independent of temperature. For this model, plots of absorbance versus temperature can be fit with four variables: ε_s, ε_u, $\Delta H°$, and $\Delta S°$, since

$$K = \exp(-\Delta H°/RT + \Delta S°/R) \tag{8-5}$$

This treatment assumes the change in absorbance is entirely due to the stacking reaction and this assumption can be checked by measuring the time dependence of the absorbance change as described in Section 8.1.2.

When single-strand stacking occurs in an oligonucleotide or polynucleotide, there is an additional consideration. A new stack can either lengthen an existing region of stacked nucleotides, or it can occur in a region that was previously completely random coil. The equilibrium constants for these two processes can be different, indicating cooperativity for stacking. In this case, a somewhat more complex model, the one-dimensional (1D) Ising model, must be used for the analysis (Zimm and Bragg, 1959; Applequist, 1963). In this model, the equilibrium constants for propagating or initiating a stacked region are denoted s and βs, respectively. Thus β for a noncooperative transition is one. As a transition becomes more cooperative, β becomes smaller. While s has the temperature dependence given above for K, β is assumed to be temperature independent. For this model, α is given by (Applequist, 1963):

$$\alpha = 0.5 + 0.5(s-1)[(1-s)^2 + 4\beta s]^{-1/2} \tag{8-6}$$

The addition of β means five parameters are required to fit data to the model. In practice, the transitions are too broad to reliably allow a five parameter fit. Simultaneous fitting of spectroscopic and calorimetric data can improve the situation, however (Freier et al., 1981).

Representative thermodynamic parameters for some dinucleoside monophosphates and polynucleotides are listed in Table 8-1. Included in the table are melting temperatures, t_m, the temperatures where half the bases are stacked ($K = 1$ and/or $\alpha = 0.5$). For any transition,

$$\Delta G° = -RT\ln K = \Delta H° - T\Delta S° \tag{8-7a}$$

Here T is the Kelvin temperature, $T = 273.15 + t$, where t is the temperature in degrees Celsius (°C). Thus, for a unimolecular transition, U ⇌ S, at the T_m, $-RT_m\ln(1) = 0 = \Delta H° - T_m\Delta S°$, so

$$T_m = \frac{\Delta H°}{\Delta S°} \quad \text{(unimolecular transition)} \tag{8-7b}$$

All temperatures in the above equations are in K. When the units of $\Delta H°$ are kilocalories per mole, they must be multiplied by 1000 if $\Delta S°$ is in entropy units (eu, cal mol^{-1} deg^{-1}). Note that round off errors often affect T_m values by several degrees. While measured T_m's are reasonably reliable, caution is required when considering values of $\Delta H°$ and $\Delta S°$ for single-strand stacking. As noted above, all involve extrapolation

Table 8.1
Thermodynamic Parameters for Single-Strand Stacking

Molecule	Reference	Method[a]	[NaCl]	β	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (eu)[c]	t_m °C
dApA	Olsthoorn et al. (1981).	CD			−7.3	−23	49
ApA	Olsthoorn et al. (1981).	CD			−7.2	−25	22
ApA	Powell et al. (1972).	A	0.01–1		−8.5	−28	26
ApU	Olsthoorn et al. (1981).	CD			−7.3	−25	22
CpC	Powell et al. (1972).	A	0.01–1		−8.5	−30	13
UpU	Simpkins and Richards (1967).	A,ORD	0.1		No stacking		< 0
Poly(A)	Filimonov and Privalov (1978).	C	0.01–0.1–		−3.0	−10	40
Poly(A)	Suurkuusk et al. (1977).	C	0.02	0.6	−8.5	−27	46
Poly(A)	Freier et al. (1981).	A,C	0.05–0.1	0.5[e]	−6.8[e]	−22	38
Poly(C)	Freier et al. (1981).	A,C	0.2	1	−9.0	−28	53
Poly(U)	Richards et al. (1963); Inners and Felsenfeld (1970).	V,S	0.0002–2		No stacking		< 0

[a]Methods: A = absorbance; C = calorimetry; CD = circular dichroism; ORD = optical rotatory dispersion; S = sedimentation; V = viscosity.
[b]Average of two values.
[c]eu = cal k^{-1} mol^{-1}

of data outside experimentally accessible temperature limits. As discussed in Section 8.1.2, kinetic results indicate the two-state model also may not be adequate for certain conditions.

Despite the uncertainties, some features of single-strand stacking are clear. The cooperativity is small, with β typically between 0.5 and 1. The salt dependence is also negligible. The sequence dependence, however, is considerable. While poly(A) and poly(C) are largely stacked at 20°C, poly (U) is a random coil. Structural studies of oligomers by NMR and ORD have provided evidence that the negligible stacking of U is not restricted to UU sequences. For example, in the sequences AUG and UGUG, the purines stack together while the U bases remain unstacked (Lee and Tinoco, 1980; van der Hoogen et al., 1988b).

1.2 Kinetics of Single Strand Stacking

1.2.1 Temperature-Jump Relaxation Spectroscopy

The kinetics of single strand stacking are very fast. They can be measured, however, by temperature-jump relaxation spectroscopy (Eigen and De Maeyer, 1963; Bernasconi, 1976; Turner, 1986). In this method, the temperature of a solution is raised quickly, which changes the equilibrium constant for any reaction that has a nonzero $\Delta H°$ (see Eq. 8-2). If the temperature-jump occurs faster than the equilibrium can adjust, then the time dependence of the approach to the new equilibrium concentrations can be followed spectroscopically. The rate constants for the reaction can be derived from this time dependence. For example, for the two-state reaction shown in Figure 8-2, the time dependence of the change in concentration of unstacked species, $\Delta[U]$, is a single exponential:

$$\Delta[U] = \Delta[U]_0 \exp(-t/\tau) \tag{8-8}$$

Here $\Delta[U]_0$ is the displacement from the new equilibrium at the time of the temperature-jump, and t is time. For a unimolecular reaction, $\tau^{-1} = k_1 + k_{-1}$, the sum of the forward and reverse rate constants. The equilibrium constant is given by $K = k_1/k_{-1}$. Thus measurements of the relaxation time, τ, and K allow determination of k_1 and k_{-1}.

1.2.2 Results

Single relaxation times have been measured for CpC, CpA, ApC, poly(C), and poly(A) (Freier et al., 1981; Pörschke, 1978; Dewey and Turner, 1979). These results are listed in Table 8-2. At Na^+ concentrations above 0.05 M, the relaxation times for poly(A) and ApA depend on the monitoring wavelength (Pörschke, 1978), which means there is more than one conformational change in these cases. Studies by NMR indicate this results because two stacked conformations exist (Kondo and Danyluk, 1976). Thus the two-state model is not appropriate, and rate constants cannot be derived easily.

All the forward rate constants for single strand stacking are about 10^7 s^{-1}. While this value is large compared to many other reactions, it is slow for a simple stacking reaction. For example, the relaxation time for stacking of the two adenines joined by

Table 8.2
Kinetic Parameters for Single Strand Stacking

Molecule	Reference	$[Na^+]$	t	τ	$k_1 \times 10^6$	E_{a1}	$\Delta S_1^\ddagger$	$k_{-1} \times 10^6$	E_{a-1}	$\Delta S_{-1}^\ddagger$
		(M)	°C	(ns)	s^{-1}	(kcal mol^{-1})	(eu)	s^{-1}	(kcal mol^{-1})	(eu)
ApA	Pörschke (1978).	1	4	50, >100	*a*			*a*		
ApC	Pörschke (1978).	1	4	42	13*			11*		
CpA	Pörschke (1978).	1	4	30	14*			19*		
CpC	Pörschke (1978).	1	4	30	19*	2.1	−19	14*	11	12
$A^9—(CH_2)_3—A^9$	Pörschke (1978).	1	4	<15						
Poly(A)	Dewey and Turner (1979).	0.05	10	174	6.2	6.2	−7.5	1.6	14	18
Poly(C)	Freier et al. (1981).	0.05	10	45	19	−0.8	−30	3.1	7.8	−2
Poly(C)	Freier et al. (1981).	1	15	67	13	1.9	−22	1.8	11	6
Poly(dA)	Dewey and Turner (1979).	0.05	10	48	22	3.3	−15	3.4	9.6	3

*Rate constants are not calculated because more than one relaxation time is observed.
*Equilibrium constants used to calculate k_1 and k_{-1} from τ were taken from Davis and Tinoco (1968) for ApC and CpA, and from Powell et al. (1972) for CpC.

a trimethylene bridge in 9,9'-trimethylenebisadenine is faster than 15 ns (Pörschke, 1978), indicating $k_1 > 7 \times 10^7 s^{-1}$. The forward rate constant for bimolecular stacking reactions of bases and of planar dye molecules is typically about $10^9 M^{-1} s^{-1}$ (Pörschke and Eggers, 1972; Dewey et al., 1979). Given the high local concentration of bases in a single-stranded polymer, the magnitude of the unimolecular rate constant would be at least as large if controlled by the same factors. These comparisons suggest the sugar–phosphate backbone somehow limits the stacking rate. Further insights into this are provided by the activation energies, E_a, and entropies, $\Delta S^\ddagger$, as derived from the temperature dependence of rate constants and the Eyring equation:

$$k = (eRT/hN) \exp(-E_a/RT + \Delta S^\ddagger/R) \tag{8-9}$$

Here e is the base for natural logarithms (2.72), h is Planck's constant, and N is Avogadro's number. For the stacking of poly(C) and poly(A), E_a is only a few kilocalories per mole (Freier et al., 1981; Dewey and Turner, 1979). Furthermore, stacking rate constants depend linearly on solvent viscosity. This finding is consistent with a diffusion controlled rate. The activation entropies, however, are about -10 to -30 eu. Thus the relatively slow rate constants are a consequence of a very ordered transition state, which could result from the necessity of constraining bonds in the backbone to single conformations before rotational diffusion to the stacked conformation. Another possible contributor to the large, unfavorable $\Delta S^\ddagger$ is specific solvation requirements for the transition state.

1.3 Interactions Determining Structure and Dynamics of Single-Strand Stacking

The results for the thermodynamics and kinetics of single-strand stacking provide insight into the interactions that determine these properties. In particular, conformational entropy favors the unstacked, random coil conformation. These unfavorable entropy effects must be overcome by bonding interactions, commonly called stacking interactions, in order to stabilize the ordered, stacked helix. Both effects are discussed below, along with some unexplained effects.

1.3.1 Conformational Entropy Effects

The covalent bonds in a dinucleoside monophosphate are shown in Figure 2-3. There are a limited number of conformations that can be adopted around each bond. For most bonds in polymers, the number of available conformations is typically three. X-ray and NMR data on nucleic acids suggest the appropriate number of available conformations for the bonds shown in Figure 2-3 are (Berman, 1981; Olson, 1982): $\alpha(1)$, $\beta(3)$, $\gamma(3)$, $\delta(2)$, $\varepsilon(3)$, $\zeta(2)$, $\chi(2)$. If the stacked state of a dinucleoside monophosphate is restricted to a single conformation, but the unstacked state can sample all available conformations for each bond, then initiation of stacking will be opposed by an unfavorable initiation conformational entropy term of

$\Delta S^\circ_{ic} = -R\ln(1 \times 3 \times 3 \times 2 \times 2 \times 3 \times 2 \times 2 \times 2) = -13$ eu. Notice that two glycosidic (χ) and two sugar (δ) bonds are included in this calculation, one for each nucleoside. If the stacked dimer in Figure 8-2 was part of a polynucleotide, then stacking an adjacent nucleotide on the pre-existing stack would require constraining rotation about one more glycosidic and one more sugar bond, rather than two of each. In this simple model, propagation of a helix is associated with a propagation conformational entropy change, ΔS°_{pc}, of -11 eu. The values measured for ΔS° and $\Delta S^\ddagger$ of stacking in dinucleoside monophosphates and in polymers range from about -10 to -30 eu (see Tables 8-1 and 8-2). Thus conformational effects can account for a large part of the measured equilibrium and activation entropy changes. They can also account for the small cooperativity observed. If the cooperativity parameter β arises only because two glycosidic and two sugar bonds are constrained for initiation and one each for propagation, then its value is predicted to be $\beta = \exp[(-R\ln 2 \times 2)/R] = 0.25$. Measured values range from about 0.5 to 1. Of course, this model may be oversimplified. For example, more than one configuration may contribute to the stacked state.

1.3.2 Stacking

The origin of stacking interactions has been the subject of much discussion. From the thermodynamic data, it is clear that stacking is driven by a favorable, modest ΔH°. This finding is consistent with noncovalent bonding due to dispersion forces as described in Section 7.1.

It has often been suggested that stacking is also driven by classical hydrophobic interactions. In this model, ordered water is released from around the bases upon stacking, and this provides a favorable entropy term. The magnitude expected for such an effect can be estimated from the dimerization of benzene in water. The ΔS° for benzene dimerization is a favorable 20 eu (Tucker et al., 1981). The ΔS° for single strand stacking, however, is unfavorable (see Table 8-1). Moreover, the magnitude is either that expected for conformational entropy effects, or is even more unfavorable. Thus the thermodynamic parameters provide no evidence that classical hydrophobic bonding is important for driving stacking.

1.3.3 Unexplained Effects

The results on single strand stacking also raise some interesting questions. Comparison of Tables 8-1 and 2-2 indicates the ΔH° values for stacking of ApA, poly(A), and deoxyadenosine monomers are similar. The ΔH° values for stacking of CpC and poly(C), however, are very different from the ΔH° for association of cytidines. Moreover, the order for strength of stacking in monomers, dimers, and polymers is different: U $\leq$ C < A versus UpU $\ll$ CpC < ApA versus poly(U) $\ll$ poly(A) < poly(C). This result suggests some new interaction in poly(C). The activation parameters for stacking in poly(C) and poly(A) are also suggestive of a difference (Table 8.2). The $E_{a,1}$ and $\Delta S^\ddagger_1$ for poly(A) are reasonable for a diffusion controlled process. For poly(C), however, $E_{a,1}$ is lower and $\Delta S^\ddagger_1$ higher than expected for diffusion control. One possibility is a special, specific solvation of poly(C).

2. DOUBLE HELIX FORMATION BY OLIGONUCLEOTIDES WITHOUT LOOPS

The main structural motif for natural nucleic acids is the double helix with Watson–Crick base pairs. Cellular DNA is almost exclusively in this form. Known structures of RNA are more than 50% double helix. Thus an understanding of the principles governing double helix formation is essential for understanding and predicting the properties of nucleic acids. Oligonucleotides provide convenient model systems for discovering these principles.

2.1 Thermodynamics of Duplex Formation

2.1.1 Methods

The methods used for measuring the thermodynamics of double helix formation by oligonucleotides are similar to those used for single-strand stacking. Double helix formation, however, is a very cooperative process, so the transitions occur in smaller temperature intervals. For example, Figure 8-4 shows a DSC curve for duplex formation

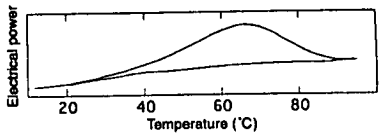

Figure 8-3
Raw data from a DSC experiment on dGCGCGC. Shown is electrical power (fed back to sample cell to maintain it at same temperature as reference cell) versus temperature. For the upper curve, the sample cell contained 1×10^{-3} *M* dGCGCGC in 1 *M* NaCl, 45 m*M* cacodylate, pH 7. For the lower curve, both sample and reference cells contained only buffer. [Reprinted with permission from Albergo, D. D., Marky, L. A., Breslauer, K. J., and Turner, D. H. (1981). *Biochemistry*, **20**, 1409–1413. Copyright ©American Chemical Society.]

by dGCGCGC (Albergo et al., 1981). The beginning and end of the transition and the baselines outside the transition region are apparent, which makes integration of the ΔC°_p curve to obtain ΔH° relatively reliable (see Eq. 8-1). To obtain ΔS°, the same data can be plotted as $\Delta C^\circ_p/T$ versus temperature and integrated:

$$\Delta S^\circ = \int_{T_1}^{T_2} \frac{\Delta C^\circ_p}{T}\, dT \qquad (8\text{-}10)$$

An advantage of calorimetry is that the thermodynamic parameters obtained in this way do not depend on a theoretical model for the transition. As long as the baselines can be determined, the thermodynamic parameters can be obtained by taking the area under the transition curve.

The temperature dependence of spectroscopic properties is also used to obtain thermodynamic parameters for duplex formation. In this case, the data must be fit to a theoretical model to derive parameters. In practice, a simple two-state model is used most often (Martin et al., 1971; Borer et al., 1974; Turner et al., 1988; Petersheim and Turner, 1983). The simplifying assumption in the two-state case is that a given strand is either maximally base paired or completely not base paired (see Fig. 8-5). This assumption corresponds to a completely cooperative transition. More general statistical mechanical treatments are discussed in Appendix 8-11.

The equations used to fit spectroscopic data to the two-state model depend on whether the oligonucleotides are self- or nonself-complementary:

$$\text{Self-complementary} \qquad 2A \rightleftharpoons A_2 \qquad K = \frac{[A_2]}{[A]^2} = \frac{\alpha}{2(1-\alpha)^2 C_T} \tag{8-11}$$

$$\text{Nonself-complementary } [A]_0 = [B]_0 = C_T/2 \qquad A + B \rightleftharpoons AB \qquad K = \frac{[AB]}{[A][B]} = \frac{2\alpha}{(1-\alpha)^2 C_T} \tag{8-12}$$

Here α is the fraction of total strand concentration, C_T, that is in duplex; and for the nonself-complementary case, it has been assumed that the total concentration of A and B strands is each $C_T/2$. If the spectroscopic property is absorbance, A, then at any temperature T

$$\text{Self-complementary} \qquad A = C_T \ell [\varepsilon_A (1-\alpha) + \varepsilon_{A_2} \alpha/2] \tag{8-13}$$

$$\text{Nonself-complementary} \qquad A = \frac{C_T \ell}{2} [(\varepsilon_A + \varepsilon_B)(1-\alpha) + \varepsilon_{AB}\alpha] \tag{8-14}$$

Here ε_A, ε_B, ε_{A_2}, and ε_{AB} are extinction coefficients for single-stranded A and B, and for duplexes A_2 and AB, respectively. If these extinction coefficients are known, then absorbance versus temperature data can be fit with Eqs. 8-11–8-14 and 8-7a to provide values of $\Delta H°$ and $\Delta S°$ for the transition. In practice, the extinction coefficients are usually temperature dependent. For example, the extinction for a single strand will vary

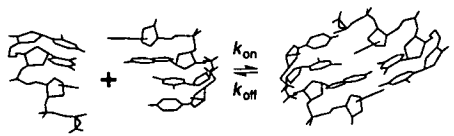

Figure 8-4
Transition from two single strands to a double helix.

with temperature because of the stacked-to-unstacked equilibrium discussed in Section 8.1. When nonself-complementary oligomers are studied, this temperature dependence can sometimes be measured independently with the individual single strands. In most cases, the temperature dependences of extinction coefficients are assumed to be linear, for example, $\varepsilon_A = m_A T + b_A$. The values for m_A and b_A are determined by fitting the absorbance versus temperature data in the regions where only single strands or only duplexes occur, or by including the linear dependence of extinction in fitting the shape of the entire duplex-to-single strand transition curve. Note that not every spectroscopic property is as simple as absorbance. For example, NMR chemical shifts cannot easily be used to derive thermodynamic properties (Pardi et al., 1981).

For a two-state transition, the concentration dependence of duplex formation provides another method for determining thermodynamic parameters. Defining the melting temperature, T_m, as the temperature at which $\alpha = 0.5$ and plugging into Eqs. 8-11 and 8-12 gives the results that at the T_m, K is $1/C_T$ for a self-complementary transition and $4/C_T$ for a nonself-complementary transition. Substituting these results into $\Delta G° = -RT \ln K = \Delta H° - T\Delta S°$, and rearranging leads to

$$\text{Bimolecular, self-complementary} \qquad \frac{1}{T_m} = \frac{R \ln C_T}{\Delta H°} + \frac{\Delta S°}{\Delta H°} \tag{8-15a}$$

$$\text{Bimolecular nonself-complementary } [A]_0 = [B]_0 = C_T/2 \qquad \frac{1}{T_m} = \frac{R \ln (C_T/4)}{\Delta H°} + \frac{\Delta S°}{\Delta H°} \tag{8-15b}$$

Thus, a plot of $1/T_m$ versus $\ln(C_T)$ should be linear, and $\Delta H°$ and $\Delta S°$ can be determined from the slope and intercept. Equations 8-15a and 8-15b are special cases of the general equations for N strands associating to form an N-mer (Marky and Breslauer, 1987):

$$N\text{-mer, Self-complementary} \qquad \frac{1}{T_m} = \frac{(N-1)R}{\Delta H°} \ln C_T + \frac{[\Delta S° - (N-1)R \ln 2 + R \ln N]}{\Delta H°} \tag{8-16a}$$

$$N\text{-mer, Nonself-complementary } [A]_0 = [B]_0 = [C]_0 = \cdots \qquad \frac{1}{T_m} = \frac{(N-1)R}{\Delta H°} \ln C_T + \frac{[\Delta S° - (N-1)R \ln 2N]}{\Delta H°} \tag{8-16b}$$

If a transition is two state, then all the methods described above should give the same thermodynamic parameters. If a transition is not two state, then the magnitude of $\Delta H°$ determined from calorimetry will be larger than that determined spectroscopically (Sturtevant, 1987). Moreover, the two spectroscopic methods may give different results, since non-two-state behavior typically affects the shape of a melting curve more than the T_m. These comparisons, therefore, provide tests for two-state behavior.

2.1.2 *Results for Duplexes without Loops*

Table 8-3 lists thermodynamic parameters determined for duplex formation by some oligonucleotides. The enthalpy and entropy changes are quite large relative to

those for single-strand stacking. In contrast to single-strand stacking, these changes also increase substantially in magnitude as the sequence is made longer (e.g. GCCGGC vs. CCGG). These trends reflect the high cooperativity of the duplex transition.

Table 8-3 contains some results for sequences that have been studied in both deoxy and ribo forms, in parallel and antiparallel strands, and in 3′–5′ and 2′–5′linked oligomers. The thermodynamic parameters are different in each case. For sequences that have been studied, duplexes with antiparallel strands are more stable than duplexes with parallel strands (Rippe and Jovin, 1992), and duplexes with 3′–5′ phosphodiester bonds are more stable than duplexes with 2′–5′phosphodiester bonds (Kierzek et al., 1992; Jin et al., 1993). For 3′–5′linked duplexes, whether a deoxy or ribo duplex is more stable depends on the sequence. Different stabilities might be expected for sequences containing AU or AT pairs. The RNA and DNA sequences with only GC pairs also have different stabilities, however. Thus the 2′-OH affects duplex stability. This 2′-OH effect is even observed if a single 2′-OH is replaced by H in an oligonucleotide (Bevilacqua and Turner, 1991). This effect may be related to the different conformational preferences for RNA and DNA (see Chapter 4). For RNA–DNA hybrid duplexes, the available data suggest interesting sequence dependence (Martin and Tinoco, 1980; Sugimoto et al., 1995). For example, at total strand concentrations of $4 \times 10^{-4} M$, the melting temperatures of $rCA_3G \cdot dCT_3G$ and $dCA_3G \cdot rCU_3G$ are 19 and less than 0°C, respectively (Martin and Tinoco, 1980). It has been suggested that the relative stabilities of RNA · RNA, DNA · DNA, and RNA · DNA duplexes may determine the positions for factor-independent termination of transcription (Martin and Tinoco, 1980).

Table 8-3 also illustrates the large dependence of stability on base pair composition. At $10^{-4} M$ oligonucleotide strands, the duplex with 9 bp formed by CA_3G and CU_3G melts at 32°C, whereas the 4 bp duplex $(GGCC)_2$ melts at 34°C. Evidently, GC pairs are more stable than AU pairs. The sequence dependence of stability depends on more than composition, however. For example, another duplex with four GC pairs, $(CGCG)_2$, melts at about 19°C at $10^{-4} M$.

Nearest neighbor models (Borer et al., 1974; Turner et al., 1988; Goldstein and Benight, 1992; Gray 1997a,b; Allawi and SantaLucia, 1997; Xia et al., 1998) provide reasonable approximations of the sequence dependence of duplex stability. One such model, the Independent Nearest Neighbor-Hydrogen Bonding or INN-HB model, assumes that the stability of a given base pair depends on the identity of the adjacent base pair and that the stability of a helix depends on these nearest neighbor interactions and the base composition of the helix as reflected by the terminal base pairs (Xia et al., 1998). For example, the nearest neighbors 5′CG3′/3′GC5′, 5′GC3′/3′CG5′, and 5′GG3′/3′CC5′ are different. Thus $(GGCC)_2$ and $(CGCG)_2$ are composed of different nearest neighbors and can have different stabilities in the model. The duplexes $(GCCGGC)_2$ and $(GGCGCC)_2$, however, both have the same nearest neighbors and same base compositions, and therefore must have identical stabilities within the model. Inspection of the experimental results in Table 8-3 for GCCGGC ($t_m = 67.2$°C) and GGCGCC ($t_m = 65.2$°C) indicates the nearest neighbor model is a reasonable approximation for this case. A study of pairs of oligonucleotides with identical nearest neighbors and base compositions but different sequences indicates the thermodynamic parameters for such oligomers are generally within about 10% of each other (Kierzek et al., 1986; Xia et al., 1998). Thus the nearest neighbor model is better than a simple composition model, but is not perfect.

Table 8.3
Thermodynamic Parameters for Duplex Formation by Oligonucleotides

		Measured[a]				Predicted[a]			
		ΔH°	ΔS°	ΔG°_{37}	t_m	ΔH°	ΔS°	ΔG°_{37}	t_m
Sequence	Reference	(kcal mol^{-1})	(eu)	(kcal mol^{-1})	(°C)	(kcal mol^{-1})	(eu)	(kcal mol^{-1})	(°C)
RNA, 3′-5′, Antiparallel Strands, 1 *M* NaCl									
CCGG	Peterkin and Turner (1983).	−34.2	−95.6	−4.6	27.1	−33.8	−95.0	−4.4	25.3
CGCG	Xia et al. (1998).	−33.3	−95.6	−3.7	19	−32.6	−93.2	−3.6	18.8
CGGC	Freier et al. (1985b).	−30.5	−83.4	−4.6	26.5	−36.8	−103.4	−4.7	29.1
GCCG	Freier et al. (1983).	−35.8	−98.1	−5.4	34.4	−38.1	−105.2	−5.4	34.9
CGCGCG	Freier et al. (1986b).	−54.5	−146.4	−9.1	57.9	−58.1	−156.8	−9.4	58.5
GCAUGC	Groebe, Cameron, Freier, Turner, and Uhlenbeck, unpublished results.	−62.3	−177.2	−7.4	45.7	−56.4	−157.2	−7.6	48.3
GCCGGC	Freier et al. (1985b).	−62.7	−166.0	−11.2	67.2	−63.6	−168.8	−11.2	66.6
GCGCGC	Freier et al. (1985b).	−66.0	−178.5	−10.6	62.1	−62.3	−167.0	−10.5	63.1
GGCGCC	Freier et al. (1985b).	−67.8	−182.0	−11.3	65.2	−63.6	−168.8	−11.2	66.6
GGCGCC	Freier et al. (1986b).	−59.8	−173.1	−5.5	31.6	−58.2	−169.5	−5.7	32.4
CA$_3$G + CU$_3$G	Nelson et al. (1981); Freier et al. (1986b).								
RNA, 2′-5′, Antiparallel Strands, 1 *M* NaCl									
GGCGGC	Kierzek et al. (1992).	−22.2	−56.2	−4.8	24.4				
CGGCCGCG	Kierzek et al. (1992).	−44.3	−121.2	−6.8	45.3				
DNA, 3′-5′, Antiparallel Strands, 1 *M* NaCl									
CGCGCG	Senior et al., (1988b).	−46.4	−122.8	−8.3	55.7	−51.2	−137.4	−8.6	55.7
GCATGC	Williams et al. (1989.	−42.4	−118	−5.9	38.3	−43.6	−121.6	−5.9	38.5
GCGCGC	Albergo et al. (1981).	−59.6	−162.7	−9.1	56.1	−50.4	−134.6	−8.7	56.5
CA$_3$G+CT$_3$G	Aboulela et al. (1985)	−68.0	−196	−7.2	40.1	−63.5	−182.5	−6.8	43.1
DNA, 3′-5′, Antiparallel Strands, 0.1 *M* NaCl									
5′A$_{10}$TA$_3$T$_3$A$_3$TAT$_3$A$_3$3′	Rippe and Jovin (1992); Remizperis (1992).	−151	−457						
3′T$_{10}$AT$_3$A$_3$T$_3$ATA$_3$5′									
5′A$_3$GA$_3$GTAGTA$_3$GTAT$_3$3′	Rippe and Jovin (1992); Remizperis (1992).	−167	−500						
3′T$_3$CGA$_3$CATCA$_3$T$_3$CATA$_3$5′									
DNA, 3′-5′, Parallel Strands, 0.1 *M* NaCl									
5′A$_{10}$TA$_3$T$_3$A$_3$TAT$_3$A$_3$3′	Rippe and Jovin (1992); Remizperis (1992).	−116	−374			−116	−354	−6.2	
5′T$_{10}$AT$_3$A$_3$T$_3$ATA$_3$3′									
5′A$_3$GA$_3$GTAGTA$_3$GTAT$_3$3′	Rippe and Jovin (1992); Remizperis (1992).	−116	−381						
3′T$_3$CT$_3$CATCA$_3$T$_3$CATA$_3$5′									

[a] Thermodynamic parameters determined from T_m^{-1} versus log C_T plots. The t_m is for $1 \times 10^{-4} M$ strand concentration.

[b] Thermodynamic parameters predicted from Table 8.4 for 1 *M* NaCl and from Rippe and Jovin (1992).Values for ΔG°_{37} calculated from the predicted ΔH° and ΔS° using $\Delta G^\circ_{37} = \Delta H^\circ - 310.15\Delta S^\circ$ may differ from the listed ΔG°_{37} due to round off errors. The t_m is for $1 \times 10^{-4} M$ strand concentration.

those for single-strand stacking. In contrast to single-strand stacking, these changes also increase substantially in magnitude as the sequence is made longer (e.g. GCCGGC vs. CCGG). These trends reflect the high cooperativity of the duplex transition.

Table 8.3 contains some results for sequences that have been studied in both deoxy and ribo forms, in parallel and antiparallel strands, and in 3′–5′and 2′–5′linked oligomers. The thermodynamic parameters are different in each case. For sequences that have been studied, duplexes with antiparallel strands are more stable than duplexes with parallel strands (Rippe and Jovin, 1992), and duplexes with 3′–5′phosphodiester bonds are more stable than duplexes with 2′–5′phosphodiester bonds (Kierzek et al., 1992; Jin et al., 1993). For 3′–5′linked duplexes, whether a deoxy or ribo duplex is more stable depends on the sequence. Different stabilities might be expected for sequences containing AU or AT pairs. The RNA and DNA sequences with only GC pairs also have different stabilities, however. Thus the 2′-OH affects duplex stability. This 2′-OH effect is even observed if a single 2′-OH is replaced by H in an oligonucleotide (Bevilacqua and Turner, 1991). This effect may be related to the different conformational preferences for RNA and DNA (see Chapter 4). For RNA–DNA hybrid duplexes, the available data suggest interesting sequence dependence (Martin and Tinoco, 1980; Sugimoto et al., 1995). For example, at total strand concentrations of 4×10^{-4} *M*, the melting temperatures of $rCA_2G \cdot dCT_2G$ and $dCA_2G \cdot rCU_2G$ are 19 and less than 0°C, respectively (Martin and Tinoco, 1980). It has been suggested that the relative stabilities of RNA · RNA, DNA · DNA, and RNA · DNA duplexes may determine the positions for factor-independent termination of transcription (Martin and Tinoco, 1980).

Table 8.3 also illustrates the large dependence of stability on base pair composition. At 10^{-4} *M* oligonucleotide strands, the duplex with 9 bp formed by CA_2G and CU_2G melts at 32°C, whereas the 4 bp duplex $(GGCC)_2$ melts at 34°C. Evidently, GC pairs are more stable than AU pairs. The sequence dependence of stability depends on more than composition, however. For example, another duplex with four GC pairs, $(CGCG)_2$, melts at about 19°C at 10^{-4} *M*.

Nearest neighbor models (Borer et al., 1974; Turner et al., 1988; Goldstein and Benight, 1992; Gray 1997a,b; Allawi and SantaLucia, 1997; Xia et al., 1998) provide reasonable approximations of the sequence dependence of duplex stability. One such model, the Independent Nearest Neighbor-Hydrogen Bonding or INN-HB model, assumes that the stability of a given base pair depends on the identity of the adjacent base pair and that the stability of a helix depends on these nearest neighbor interactions and the base composition of the helix as reflected by the terminal base pairs (Xia et al., 1998). For example, the nearest neighbors 5′CG3′/3′GC5′, 5′GC3′/3′CG5′, and 5′GG3′/3′CC5′ are different. Thus $(GGCC)_2$ and $(CGCG)_2$ are composed of different nearest neighbors and can have different stabilities in the model. The duplexes $(GCCGGC)_2$ and $(GGCCGC)_2$, however, both have the same nearest neighbors and same base compositions, and therefore must have identical stabilities within the model. Inspection of the experimental results in Table 8.3 for GCCGGC (t_m = 67.2°C) and GGCGCC (t_m = 65.2°C) indicates the nearest neighbor model is a reasonable approximation for this case. A study of pairs of oligonucleotides with identical nearest neighbors and base compositions but different sequences indicates the thermodynamic parameters for such oligomers are generally within about 10% of each other (Kierzek et al., 1986; Xia et al., 1998). Thus the nearest neighbor model is better than a simple composition model, but is not perfect.

Table 8.3
Thermodynamic Parameters for Duplex Formation by Oligonucleotides

		Measured[a]				Predicted[b]			
Sequence	Reference	ΔH° (kcal mol^{-1})	ΔS° (eu)	ΔG°_{37} (kcal mol^{-1})	t_m (°C)	ΔH° (kcal mol^{-1})	ΔS° (eu)	ΔG°_{37} (kcal mol^{-1})	t_m (°C)
RNA, 3′-5′, Antiparallel Strands, 1 *M* NaCl									
CCGG	Petersheim and Turner (1983).	−34.2	−95.6	−4.6	27.1	−33.8	−95.0	−4.4	25.3
CGCG	Xia et al. (1998).	−33.3	−95.6	−3.7	19	−32.6	−93.2	−3.6	18.8
GCGC	Freier et al. (1985b).	−30.5	−83.4	−4.6	26.5	−36.8	−103.4	−4.7	29.1
GGCC	Freier et al. (1983).	−35.8	−98.1	−5.4	34.4	−38.1	−105.2	−5.4	34.9
CGCGCG	Freier et al. (1986b).	−54.5	−146.4	−9.1	57.9	−58.1	−156.8	−9.4	58.5
GCAUGC	Groebe, Cameron, Freier, Turner, and Uhlenbeck, unpublished results.	−62.3	−177.2	−7.4	45.7	−56.4	−157.2	−7.6	48.3
GCCGGC	Freier et al. (1985b).	−62.7	−166.0	−11.2	67.2	−63.6	−168.8	−11.2	66.6
GCGCGC	Freier et al. (1985b).	−66.0	−178.5	−10.6	62.1	−62.3	−167.0	−10.5	63.1
GGCGCC	Freier et al. (1986b).	−67.8	−182.0	−11.3	65.2	−63.6	−168.8	−11.2	66.6
$CA_2G + CU_2G$	Nelsen et al. (1981); Freier et al. (1986b).	−59.8	−175.1	−5.5	31.6	−58.2	−169.5	−5.7	32.4
RNA, 2′-5′, Antiparallel Strands, 1 *M* NaCl									
GCGCGC	Kierzek et al. (1992).	−22.2	−56.2	−4.8	24.4				
CGGCGCCG	Kierzek et al. (1992).	−44.3	−121.2	−6.8	45.3				
DNA, 3′-5′, Antiparallel Strands, 1 *M* NaCl									
CGCGCG	Senior et al., (1988b).	−46.4	−122.8	−8.3	55.7	−51.2	−137.4	−8.6	55.7
GCATGC	Williams et al. (1989.	−42.4	−118	−5.9	38.3	−43.6	−121.6	−5.9	38.5
GCGCGC	Albergo et al. (1981).	−59.6	−162.7	−9.1	56.1	−50.4	−134.6	−8.7	56.5
CA_2G+CT_2G	Aboulela et al. (1985)	−68.0	−196	−7.2	40.1	−63.5	−182.5	−6.8	43.1
DNA, 3′-5′, Antiparallel Strands, 0.1 M NaCl									
$5'A_{10}TA_2T_2A_3TAT_3 3'$ $3'T_{10}AT_2A_4T_3ATA_3 5'$	Rippe and Jovin (1992); Rentzeperis (1992).	−151	−457						
$5'A_3GA_3GTAGT_4A_2GTAT_3 3'$ $3'T_3CG_2CATCA_4T_2CATA_3 5'$	Rippe and Jovin (1992); Rentzeperis (1992).	−167	−500						
DNA, 3′-5′, Parallel Strands, 0.1 *M* NaCl									
$5'A_{10}TA_2T_2A_3TAT_3 3'$ $5'T_{10}AT_2A_4T_3ATA_3 3'$	Rippe and Jovin (1992); Rentzeperis (1992).	−116	−374			−116	−354	−6.2	
$5'A_3GA_3GTAGT_4A_2GTAT_3 3'$ $5'T_3CT_3CATCA_4T_2CATA_3 3'$	Rippe and Jovin (1992); Rentzeperis (1992).	−116	−381						

[a]Thermodynamic parameters determined from T_m^{-1} versus log C_T plots. The t_m is for 1×10^{-4} *M* strand concentration.

[b]Thermodynamic parameters predicted from Table 8.4 for 1 *M* NaCl and from Rippe and Jovin (1992). Values for ΔG°_{37} calculated from the predicted ΔH° and ΔS° using $\Delta G^\circ_{37} = \Delta H^\circ - 310.15\Delta S^\circ$ may differ from the listed ΔG°_{37} due to round off errors. The t_m is for 1×10^{-4} *M* strand concentration.

Table 8.4
Thermodynamic Parameters for Helix Initiation and Propagation in 1*M* NaCl

Propagation Sequence	RNA[a] ΔH° (kcal mol^{-1})	RNA[a] ΔS° (eu)	RNA[a] ΔG°_{37} (kcal mol^{-1})	DNA[b] ΔH° (kcal mol^{-1})	DNA[b] ΔS° (eu)	DNA[b] ΔG°_{37} (kcal mol^{-1})	RNA/DNA[c] ΔH° (kcal mol^{-1})	RNA/DNA[c] ΔS° (eu)	RNA/DNA[c] ΔG°_{37} (kcal mol^{-1})
RNA DNA									
→ GC / CG ←	−14.88	−36.9	−3.42	−9.8	−24.4	−2.24	−8.0	−17.1	−2.7
→ GG / CC ←	−13.39	−32.7	−3.26	−8.0	−19.9	−1.84	−12.8	−31.9	−2.9
							−9.3	−23.2	−2.1
→ CG / GC ←	−10.64	−26.7	−2.36	−10.6	−27.2	−2.17	−16.3	−47.1	−1.7
→ GA / CU ← → GA / CT ←	−12.44	−32.5	−2.35	−8.2	−22.2	−1.30	−5.5	−13.5	−1.3
							−8.6	−22.9	−1.5
→ GU / CA ← → GT / CA ←	−11.40	−29.5	−2.24	−8.4	−22.4	−1.44	−7.8	−21.6	−1.1
							−5.9	−12.3	−2.1
→ CA / GU ← → CA / GT ←	−10.44	−26.9	−2.11	−8.5	−22.7	−1.45	−9.0	−26.1	−0.9
							−10.4	−28.4	−1.6
→ CU / GA ← → CT / GA ←	−10.48	−27.1	−2.08	−7.8	−21.0	−1.28	−7.0	−19.7	−0.9
							−9.1	−23.5	−1.8
→ UA / AU ← → TA / AT ←	−7.69	−20.5	−1.33	−7.2	−21.3	−0.58	−7.8	−23.2	−0.6
→ AU / UA ← → AT / TA ←	−9.38	−26.7	−1.10	−7.2	−20.4	−0.88	−8.3	−23.9	−0.9
→ AA / UU ← → AA / TT ←	−6.82	−19.0	−0.93	−7.9	−22.2	−1.00	−7.8	−21.9	−1.0
							−11.5	−36.4	−0.2
For Bimolecular Associations									
Initiation	3.61	−1.5	4.09	0.2	−5.6	1.96	1.9	−3.9	3.1
Each terminal AU or AT	3.72	10.5	0.45	2.2	6.9	0.05			
Symmetry correction (self-complementary)	0	−1.4	0.43	0	−1.4	0.43			
Symmetry correction (nonself-complementary)	0	0	0	0	0	0			

[a]Xia et al., (1998).
[b]Allawi and SantaLucia (1997).
[c]Sugimoto et al., (1995). For RNA/DNA hybrids with two sets of parameters, the top set corresponds to the top strand as RNA and the bottom set corresponds to the bottom strand as RNA. For example, $\Delta H^\circ_{\substack{5'rGG3'\\3'dCC5'}} = -12.8$ kcal mol^{-1} and $\Delta H^\circ_{\substack{5'dGG3'\\3'rCC5'}} = -9.3$ kcal mol^{-1}. Alternative analyses of the RNA/DNA data base have also been presented (Gray, 1997b).

Application of the INN-HB nearest neighbor model to RNA and DNA duplexes containing only Watson–Crick pairs requires determining thermodynamic parameters for 10 nearest neighbors, helix initiation, and helix termination by AU or AT for each case. For RNA/DNA hybrids, more parameters are required (Sugimoto et al., 1995; Gray, 1997a,b). The parameters have been determined by fitting to optical melting data for oligonucleotides (Xia et al., 1998; Allawi and SantaLucia, 1997; Sugimoto et al., 1995; Gray, 1997b). Some results are listed in Table 8-4, and comparisons of predictions from the model with measurements are shown in Table 8-3. For DNA, calorimetric data for both oligomers and polymers have also been fit to a similar model (Breslauer et al., 1986). SantaLucia (1998) and Owczarzy et al. (1997) compare the thermodynamic parameters for DNA obtained by various methods. A nearest neighbor analysis has also been done for the available data on parallel stranded double helixes (Rippe and Jovin, 1992).

In deriving the parameters in Table 8-4, it was necessary to consider the fact that duplexes formed by self-complementary oligomers have a twofold rotational symmetry, whereas single-strands and nonself-complementary oligomers do not. To correct for this symmetry difference, ΔS° for self-complementary oligomers must be reduced by $R\ln 2 = 1.4$ eu. Including this correction and simply summing up appropriate parameters allows prediction of thermodynamic parameters for the melting of any RNA or DNA duplex containing only Watson–Crick base pairs. The melting temperature in degrees Celsius, t_m, for any concentration can then be calculated from a rearrangement of Eqs. 8-15a and 8-15b:

$$\text{Bimolecular, self-complementary} \quad t_m = \frac{\Delta H^\circ}{\Delta S^\circ + R\ln(C_T)} - 273.15 \tag{8-17}$$

$$\text{Bimolecular nonself-complementary } ([A]_0 = [B]_0 = C_T/2) \quad t_m = \frac{\Delta H^\circ}{\Delta S^\circ + R\ln(C_T/4)} - 273.15 \tag{8-18a}$$

Equation 8-18a holds if the concentrations of the two nonself-complementary strands are equal. In many applications, including probing with oligonucleotides for complementary sequences in polynucleotides, one strand is in large excess. In this case, when t_m is defined as the temperature where half of the less concentrated sequence is bound:

$$\text{Bimolecular nonself-complementary } [B]_0 \gg [A]_0 \quad t_m = \frac{\Delta H^\circ}{\Delta S^\circ + R\ln(C_B - 0.5C_A)} - 273.15 \tag{8-18b}$$

Equations 8-17 and 8-18a can be combined into one by including all the changes between self- and nonself-complementary oligomers in a constant, A:

$$t_m = \frac{\Delta H^\circ}{A + \Delta S^\circ_{NN} + R\ln(C_T)} - 273.15 \tag{8-19}$$

Here, ΔS°_{NN} is the entropy change without any symmetry term, C_T is always the total strand concentration, and A is -1.4 and -2.8 eu for self- and nonself-complementary oligomers, respectively. Note that in Eqs. 8-17–8-19, if the units for ΔH° are in kilocalories per mole, they must be multiplied by 1000 if ΔS° is in entropy units. Sample calculations are shown in Figure 8-6.

The results in Table 8-4 allow prediction of the thermodynamic properties of fully paired duplexes. Many nucleic acid associations, however, include additional unpaired nucleotides on the ends of double helixes. Two examples are the associations of tRNA and mRNA and of hybrid probes with DNA or RNA. For RNA, these unpaired nucleotides or "dangling ends" add stability to the double helix in a manner that is very sequence dependent (Turner et al., 1988). For example, at 10^{-4} M strands, the melting temperatures of $(GGCC)_2$, $(AGGCC)_2$ and $(GGCCA)_2$ are 34.4, 38.1, and 57.9°C, respectively (Freier et al., 1983b, 1985a). The effects of at least the first unpaired nucleotide adjacent to a helix can be approximated by a nearest neighbor model. The available parameters are listed in Table 8-5 (Turner et al., 1988). One striking observation for RNA oligomers in 1 M NaCl is that all the free energy increments for terminal unpaired nucleotides on the 5′ side of helixes are very similar and small, averaging -0.2 kcal mol^{-1}. Similar free energy increments are measured for addition of a 5′-phosphate suggesting that 5′-dangling ends interact little with the adjacent helix (Freier et al., 1985). Stability increments for 3′ dangling ends are sequence dependent, however, ranging from -0.1 to -1.7 kcal mol^{-1}. Thus some 3′ dangling ends can stabilize a helix more than some base pairs (cf. Tables 8-5 and 8-4), indicating a strong interaction. This difference between 5′ and 3′ stacking can be rationalized from structural considerations. Figure 8-7 shows stereoviews of $(AGGCC)_2$ and $(GGCCA)_2$ in A- form geometry. Whereas the 3′-dangling A of $(GGCCA)_2$ stacks directly on the opposite strand G of the adjacent base pair, the 5′-dangling A of $(AGGCC)_2$ is not close to the opposite strand C of the adjacent base pair. Thus it is not surprising that interactions of the 3′A with the opposite strand help hold the duplex together, whereas interactions of the 5′A with the opposite strand are negligible. Fewer data are available for DNA, but it appears that 5′-dangling ends will add more stability than 3′-dangling ends (Senior et al., 1988b; Mellema et al., 1984).

A 5′ and 3′ dangling end opposite each other is called a terminal mismatch. Free energy increments associated with terminal mismatches are listed in Table 8-6 (Freier et al., 1986a; Hickey and Turner, 1985a; Sugimoto et al., 1987b; Serra et al., 1994). For pyrimidine–pyrimidine and CA mismatches, these stability increments are essentially the sum of the increments for the constituent dangling ends. For purine–purine mismatches, the increment may be less. Thus there is no thermodynamic evidence for interactions between the bases in most terminal mismatches. The exceptions are terminal GU mismatches. In certain cases, the pairing of a terminal G and U provides more stability than the sum of the equivalent dangling ends. This finding is consistent with hydrogen-bond formation within GU mismatches as suggested by Crick (1966) in the "wobble" hypothesis (see Chapter 2).

2.2 Interactions Determining Stability of the Double Helix

Conformational entropy and stacking interactions present in single-stranded nucleic acids are also present in duplexes. In addition, hydrogen bonds are formed between base pairs, and duplex formation leads to increased condensation of counterions around

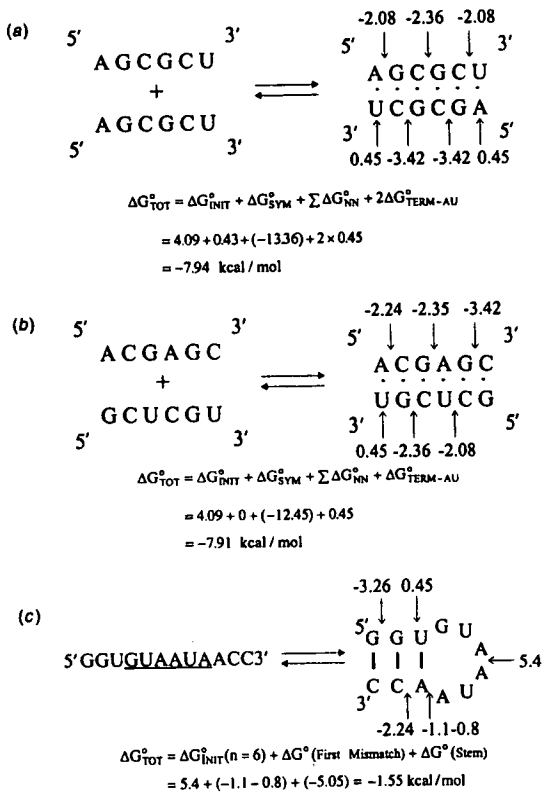

Figure 8-5
Predicting stability with the nearest neighbor model of Xia et al. (1998). (*a*) Calculation of free energy change for duplex formation by a self-complementary sequence. (*b*) Calculation for nonself-complementary sequence. Calculations of $\Delta H°$ and $\Delta S°$ are similar to those for $\Delta G°$, except there is no symmetry term for $\Delta H°$. Thus for the top sequence, $\Delta H° = 3.61 + 2(-10.48) + 2(-14.88) + (-10.64) + 2(3.72) = -50.31$ kcal mol^{-1}, $\Delta S° = -1.5 - 1.4 + 2(-27.1) + 2(-36.9) + (-26.7) + 2(10.5) = -136.6$ eu, and t_m at 10^{-4} *M* strands = 324.8 K = 51.6°C. For the sequence in (*b*), $\Delta H° = 3.61 + (-11.40) + (-10.64) + (-12.44) + (-10.48) + (-14.88) + 3.72 = -52.51$ kcal mol^{-1}, $\Delta S° = -1.5 + (-29.5) + (-26.7) + (-32.5) + (-27.1) + (-36.9) + 10.5 = -143.7$ eu, and t_m at 10^{-4} *M* strands = 318.7 K = 45.6°C. Calculations for intramolecular folding are illustrated in (*c*) and Figure 8-16. For intramolecular folding, $\Delta G°_{INIT}$ is replaced by $\Delta G°_{loops}$, and there are no symmetry terms. Additional examples are given elsewhere (Xia et al., 1998).

Table 8.5
Thermodynamic Parameters for Unpaired Terminal Nucleotides in RNA at 1 *M* NaCl[a]

Propagation	X=A			X=C			X=G			X=U		
Sequence	$\Delta H°$	$\Delta S°$	$\Delta G°_{37}$	$\Delta H°$	$\Delta S°$	$\Delta G°_{37}$	$\Delta H°$	$\Delta S°$	$\Delta G°_{37}$	$\Delta H°$	$\Delta S°$	$\Delta G°_{37}$
3′ Unpaired Nucleotides												
↑AX U ↓	−4.9	−13.2	−0.8	−0.9	−1.2	−0.5	−5.5	−15.0	−0.8	−2.3	−5.4	−0.6
↑CX G ↓	−9.0	−23.4	−1.7	−4.1	−10.7	−0.8	−8.6	−22.2	−1.7	−7.5	−20.4	−1.2
↑GX C ↓	−7.4	−20.0	−1.1	−2.8	−7.9	−0.4	−6.4	−16.6	−1.3	−3.6	−9.7	−0.6
↑UX A ↓	−5.7	−16.4	−0.7	−0.7	−1.8	−0.1	−5.8	−16.4	−0.7	−2.2	−6.8	−0.1
5′ Unpaired Nucleotides												
↑XA U ↓	1.6	6.1	−0.3	2.2	7.9	−0.3	0.7	3.4	−0.4	3.1	10.6	−0.2
↑XC G ↓	−2.4	−6.0	−0.5	3.3	11.8	−0.3	0.8	3.4	−0.2	−1.4	−4.3	−0.1
↑XG C ↓	−1.6	−4.5	−0.2	0.7	3.1	−0.3	−4.6	−14.8	0.0	−0.4	−1.2	0.0
↑XU A ↓	−0.5	−0.7	−0.3	6.9	22.8	−0.1	(0.6)	(2.7)	(−0.2)	(0.6)	(2.7)	(−0.2)

[a]Turner et al., (1988). The parameters $\Delta H°$ and $\Delta G°$ (in kcal mol^{-1}); $\Delta S°$ (in eu). Values in parentheses are estimated.

Table 8.6
Thermodynamic Parameters for Terminal Mismatches in RNA[a]

ΔG°_{37} (kcal mol^{-1})

X/Y	A	C (5′ GX / 3′ CY)	G	U
A	−1.1	(−1.5)	−1.3	(−0.5)
C	−1.1	(−0.7)	−1.4	(−0.7)
G	−1.6	(−1.0)		
U				

X/Y	A	C (5′ CX / 3′ GY)	G	U
A	−1.5	−1.5	−1.4	−0.8
C	(−1.0)	(−1.1)	−1.6	−1.2
G	−1.4	−1.4		
U				

X/Y	A	C (5′ AX / 3′ UY)	G	U
A	−0.8	(−1.0)	(−0.8)	(−0.7)
C	−0.6	(−0.7)	(−0.8)	(−0.8)
G	−0.8	(−0.8)		
U				

X/Y	A	C (5′ UX / 3′ AY)	G	U
A	−1.0	−0.8	−1.1	−0.5
C	−0.7	−0.6	−1.2	−0.5
G	−1.1	−0.6		
U				

ΔH° (kcal mol^{-1})

X/Y	A	C (5′ GX / 3′ CY)	G	U
A	−5.2	(−4.0)	−5.6	(−4.2)
C	−7.2	(+0.5)	−6.2	(−5.0)
G	−7.1	(−0.3)		
U				

X/Y	A	C (5′ CX / 3′ GY)	G	U
A	−9.1	−5.6	−5.6	−2.7
C	(−5.7)	(−3.4)	−9.2	−8.6
G	−8.2	−5.3		
U				

X/Y	A	C (5′ AX / 3′ UY)	G	U
A	−3.9	(+2.0)	(−3.5)	(−0.3)
C	−2.3	(+6.0)	(−3.5)	(−1.7)
G	−3.1	(+4.6)		
U				

X/Y	A	C (5′ UX / 3′ AY)	G	U
A	−4.0	−6.3	−8.9	−1.8
C	−4.3	−5.1	−8.9	+1.4
G	−3.8	−1.4		
U				

ΔS° (eu)

X/Y	A	C (5′ GX / 3′ CY)	G	U
A	−13.2	(−8.2)	−13.9	(−12.2)
C	−19.6	(+3.9)	−15.1	(−14.0)
G	−17.8	(+2.1)		
U				

X/Y	A	C (5′ CX / 3′ GY)	G	U
A	−24.5	−13.5	−13.4	−6.3
C	(−15.2)	(−7.6)	−24.6	−23.9
G	−21.8	−12.6		
U				

X/Y	A	C (5′ AX / 3′ UY)	G	U
A	−10.2	(+9.6)	(−8.7)	(+1.5)
C	−5.3	(+17.4)	(−8.7)	(−2.7)
G	−7.3			
U				

X/Y	A	C (5′ UX / 3′ AY)	G	U
A	−9.7	−17.7	−25.2	−4.2
C	−11.6	−14.6	−25.0	+6.0
G	−8.5	−2.5		
U				

[a]Freier et al. (1986a); Hickey and Turner, (1985); Sugimoto et al. (1987b); Serra et al. (1994). Values in parentheses are estimated.

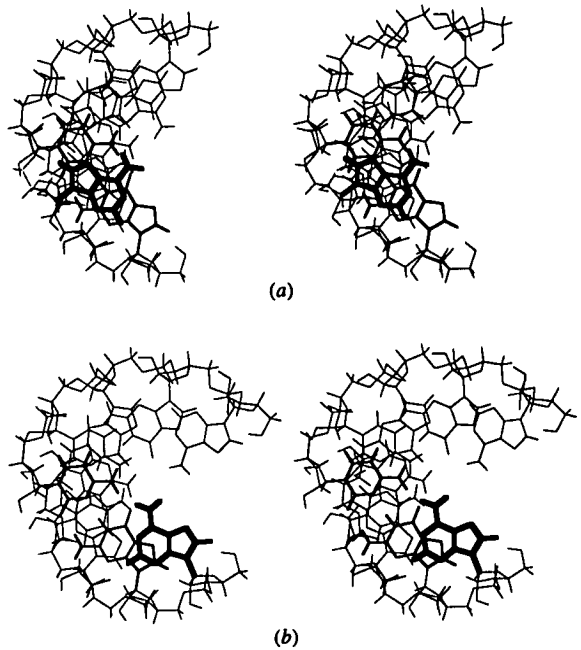

Figure 8-6
Stereoviews of $(GGCCA)_2$ (*a*) and $(AGGCC)_2$ (*b*) in A-form geometry. The terminal A closest to the reader is in boldface.

the backbone. All these factors affect the stability of the duplex. Solvent effects may also play a role. While exact partitioning of these effects on stability is not yet possible, a qualitative picture is emerging, and is discussed below.

2.2.1 *Conformational Entropy*

In Section 1.3.1, the conformational entropy associated with propagating a single-strand stacked helix by one additional nucleotide was estimated as −11 eu. Since propagation of a double helix by an additional base pair requires limiting the conformations accessible to two nucleotides, the conformational entropy associated with this process is estimated as $2 \times -11 = -22$ eu. Inspection of Table 8-4 indicates that 3/4 of the measured ΔS° values for duplex propagation are within 6 eu of this theoretical

value. As with single strands, this suggests conformational entropy effects account for a large part of the observed ΔS°.

Conformational entropy effects may also account for the stabilization of RNA duplexes by 5′-dangling ends. As noted above, stabilization from a 5′-dangling end is about the same as from a 5′-phosphate. It has been suggested by Sundaralingam that addition of a 5′-phosphate restricts the conformations of the ribose group (Sundaralingam, 1969, 1973). This restriction could provide the small stabilization that is observed.

2.2.2 *Stacking*

An empirical indication of the contributions of stacking to duplex stability is provided by the stability increments for 3′-dangling ends in Table 8-5. Comparison of some selected increments with those for base pairs is provided in Figure 8-8. This comparison is deceptive when considering the favorable attractive forces of stacking interactions, however, because the favorable free energy increment for stacking of most 3′-dangling ends includes the unfavorable conformational entropy associated with that stacking. This unfavorable component of the stacking ΔG° has been empirically estimated at about 1.9 kcal mol^{-1} at 37°C (Freier et al., 1986c), somewhat less than the value of $-T\Delta S^\circ = -(-11 \times 310) = 3.4$ kcal mol^{-1} expected from the theoretical considerations given above. When the empirical estimate of conformational effects is factored out, the favorable stacking interactions are estimated to be as large as $-1.7 - 1.9 = -3.6$ kcal mol^{-1}. The magnitudes of these favorable interactions are very sequence dependent. In fact, the negligible ΔG° values measured for $\overset{\rightarrow}{\underset{A}{UU}}$ and $\overset{\rightarrow}{\underset{A}{UC}}$ may indicate no stacking at all for these sequences. This sequence dependence is consistent with the idea that electronic interactions between the bases are responsible for stacking.

2.2.3 *Hydrogen Bonding*

Inspection of Figure 8-8 indicates the stability increments for some base pairs are much larger than the increments for stacking of the constituent nucleotides. This suggests pairing between the bases contributes an additional interaction (Freier et al., 1986c; 1985a). Consistent with this are results with DNA oligomers showing that substitution of weakly hydrogen bonding difluorotoluene for T in an AT pair decreases duplex stability by 3.6 kcal mol^{-1} at 25°C (Moran et al., 1997). Presumably, the effect of hydrogen bonds is due to the difference between hydrogen bonding in a base pair and hydrogen bonding of the separated bases with water. Specific solvation effects could also be important. Estimates for the contributions of differential hydrogen bonding can be made by taking stability increments for base pairs, subtracting stability increments for stacking of the constituent nucleotides, and correcting for conformational entropy effects. For example (Fig. 8-8), in $(GCCGGC)_2$, the ΔG°_{37} for adding each terminal GC pair is $[\Delta G^\circ_{37}(GCCGGC) - \Delta G^\circ_{37}(CCGG)]/2 = -3.3$ kcal mol^{-1}. Stacking of each 3′-dangling end in $(CCGGC)_2$ is associated with a ΔG°_{37} of

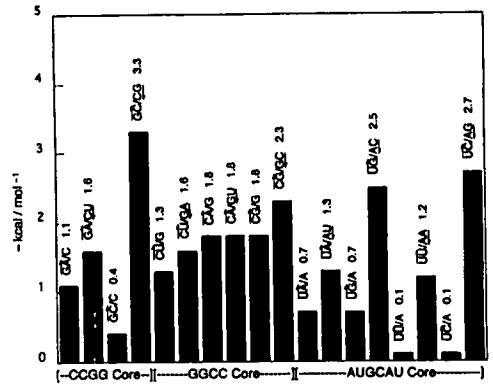

Figure 8-7
Free energy increments associated with duplex formation at 37°C from adding a terminal base pair or 3′ unpaired terminal nucleotide to CCGG, GGCC, or AUGCAU cores. For example, the first bar on the left is ΔG°_{37} for adding a 3′ unpaired A adjacent to a GC pair, ΔG°_{37} $\left(\begin{smallmatrix}5'GA3'\\ C\end{smallmatrix}\right) = 1/2\,[\Delta G^\circ_{37}(CCGGA) - \Delta G^\circ_{37}(CCGG)] = -1.1$ kcal mol^{-1}. Note that the free energy increments for adding 5′ unpaired terminal nucleotides are not shown but average only -0.2 kcal mol^{-1}.

$[\Delta G^\circ_{37}(CCGGC) - \Delta G^\circ_{37}(CCGG)]/2 = -0.4$ kcal mol^{-1}. The effect of the 5′-terminal G is $[\Delta G^\circ_{37}(GCCGG) - \Delta G^\circ_{37}(CCGG)]/2 = -0.2$ kcal mol^{-1}. (This is no more than the contribution from adding a 5′ terminal phosphate suggesting an unpaired 5′-terminal G is not stacked.) Constraining the 5′-terminal G in a base pair, however, requires overcoming the estimated 1.9 kcal mol^{-1} of conformational free energy. Thus the free energy gained from pairing the terminal G and C is estimated as $\Delta G^\circ_{37,p} = -3.3 - (-0.4 - 0.2 + 1.9) = -4.6$ kcal mol^{-1}. Since this pairing involves three hydrogen bonds, the estimated free energy increment per hydrogen bond is -1.5 kcal mol^{-1} hydrogen bond. Similar calculations on other sequences give ΔG° values for a hydrogen bond that range from -0.5 to -1.5 kcal mol^{-1} hydrogen bond (Freier et al., 1986c; Turner et al., 1987). It has been suggested that this range is due to a sequence dependent competition between hydrogen bonding and stacking (Turner et al., 1987).

Another empirical estimate for the contributions of hydrogen bonds is provided by comparing duplex stabilities for sequences with different numbers of hydrogen bonds (Freier et al., 1986c; Turner et al., 1987). Such comparisons must also consider changes in stacking interactions that may result from changing the number of hydrogen bonding groups on the bases. Inspection of Table 8-5, however, indicates that at least for stacking of 3′-terminal unpaired bases, there is little effect of changing hydrogen-bonding

groups. Both A and G have very similar thermodynamic parameters for 3′ stacking, and U and C are also similar. Thus any large changes in stacking energies between AU and GC pairs are probably a second-order effect resulting from redistribution of electrons on pairing. Even these effects can be minimized by comparing stabilities of base pairs with single changes in structure. For example, removing the 2-amino group from G in a GC pair leaves an IC pair with two hydrogen bonds. Table 8-7 contains several comparisons of oligomers with similar sequences, but different numbers of hydrogen bonds. The stability increments per hydrogen bond range from -0.5 to -2 kcal mol^{-1}.

2.2.4 Counterion Condensation

Repulsions of negatively charged backbone phosphates destabilize double helixes. Condensation of positively charged counterions around the helix favors duplex stability by neutralizing much of the negative charge on the phosphates. Experimental and theoretical aspects of this effect are discussed in Section 8.5, and in Chapter 11, respectively.

2.2.5 Solvent

By analogy to proteins, classical hydrophobic interactions are often invoked as a source of stability for double helixes. There is no evidence, however, to support the view that classical hydrophobic interactions stabilize duplexes or single-strand stacking. As discussed above, in most cases, the $\Delta S°$ of duplex formation is similar to or more unfavorable than expected from conformational terms. There is no indication of the large, positive $\Delta S°$ expected for classical hydrophobic interactions. Hydrophobic interactions are also associated with a large change in heat capacity, $\Delta C_p°$. Values of $\Delta C_p°$ reported for single- to double-strand transitions of nucleic acids, however, are either modest or zero (Suurkuusk et al., 1977; Freier et al., 1983b, 1985a, 1986a,c; Kierzek et al., 1986; Breslauer et al., 1986). These generalizations even hold for duplexes containing thymine. Since thymine has a methyl group, some hydrophobic interaction might be expected.

One solvent effect that may be important for helix stability is specific solvation. This topic is discussed in detail in Chapters 4, 7, and 11. Such specific solvation could explain the excess unfavorable $\Delta S°$ observed for some helix transitions. The quantitative consequences of such solvation have not been determined, however.

2.3 Kinetics of Duplex Formation

2.3.1 Experimental Measurements

The kinetics of duplex formation by oligonucleotides have been measured primarily with the temperature-jump method. For the reaction shown in Figure 8-5, the time dependence of the change in concentration of single strands has the same form as Eq. 8-8. For duplex formation, the relaxation time, τ, is given by

$$\text{Self-complementary} \qquad \tau^{-1} = 4k_{on}[A] + k_{off} \tag{8-20}$$

$$\text{Nonself-complementary} \qquad \tau^{-1} = k_{on}([A] + [B]) + k_{off} \tag{8-21}$$

Table 8.7
Comparison of Stabilities of Duplexes with Different Numbers of Hydrogen Bonds[a]

Reference Duplex	Duplex with Fewer H Bonds	Reference	Differences in No. of H Bonds	$\Delta\Delta t_m$ at 10^{-4} M(°C)	$\Delta\Delta G°_{37}$/H Bond (kcal mol^{-1} H Bond)
GCCGGC	ICCGGC	Turner et al. (1987)	2	15.9	−1.6
CGGCCG	CGGCCI	Turner et al. (1987)	2	8.5	−0.7
dCA$_3$GA$_3$G+dCT$_3$CT$_3$G	dCA$_3$IA$_3$G+dCT$_3$CT$_3$G	Martin et al. (1985)	1	2.8	−0.5
dCA$_3$CA$_3$G +dCT$_3$GT$_3$G	CA$_3$CA$_3$G +dCT$_3$IT$_3$G	Aboul-ela et al. (1985)	1	8.9	−1.8
dG$_3$A$_3$GCT$_3$C$_3$	dG$_3$A$_3$ICT$_3$C$_3$	Kawase et al. (1986)	2		−1.3
dCGTA′CG[a]	dCGTACG	Gaffney et al. (1984)	2		−0.5
CGGCCG	CAGCUG	Freier et al. (1986b)	2	20.1	−1.6
GCGCGC	GUGCAC	Freier et al. (1986b)	2	14.5	−1.5
GCGCGC	GCAUGC	Freier et al. (1986b)	2	16.4	−1.6
CGCGCG	UGCGCA	Freier et al. (1986b)	2	4.8	−0.5
GAUGCAUC	AAUGCAUU	Freier et al. (1986b)	2	12.2	−1.5
AUGCGCAU	AUACGUAU	Freier et al. (1986b)	2	18.4	−1.8
AUGCGCAU	AUGUACAU	Freier et al. (1986b)	2	18.7	−1.9

[a] I is inosine; A′ is 2-aminoadenine.

Here [A] and [B] are the equilibrium concentrations of the single strands at the higher temperature; k_{on} and k_{off} are the forward and reverse rate constants at the higher temperature as illustrated in Figure 8-5. Thus a plot of τ^{-1} versus single-strand concentration has an intercept of k_{off} and a slope of $4k_{on}$ or k_{on} depending on whether the sequence is self- or nonself-complementary. The activation energies for k_{on} and k_{off} can be determined from measurements as a function of temperature (see Eq. 8-9).

Typical results for the kinetics of duplex formation are compiled in Table 8-8. The forward rates, k_{on}, range from about 10^5–$10^7\ M^{-1}s^{-1}$, and depend on both sequence and salt concentration. Higher salt concentrations increase the forward rate as expected for association of two like-charged species. The forward rate, however, is slower than the diffusion limit. Moreover, in several cases, the activation energy for k_{on} is negative or zero. That is, k_{on} can decrease with increasing temperature. Any elementary reaction step must have a positive activation energy. For example, diffusion controlled reactions in solution have activation energies of about 4 kcal mol^{-1}. Thus the mechanism of duplex formation must be more complicated than indicated in Figure 8-5.

Figure 8-9 illustrates a mechanism consistent with the kinetic experiments (Pörschke and Eigen, 1971; Craig et al., 1971; Williams et al., 1989; Wetmur and Davidson, 1968; Manning, 1976). In this mechanism, the rate-determining step is formation of a nucleus containing a small number of base pairs. This nucleus adds an additional base pair faster than it dissociates. Thus the double helix can "zip up" after formation of the nucleus. The stability of the intermediate preceding the nucleus, however, is temperature dependent. Like any double helix, its stability decreases as temperature increases. Thus the concentration of the intermediate preceding the nucleus decreases as temperature increases, and this can lead to a decrease in k_{on}. The measured activation energy for k_{on} depends on the ΔH° associated with formation of the intermediate preceding the nucleus and the activation energy for forming the last base pair in the nucleus (see Fig. 8-10): $E_{A,on} = \Delta H^\circ_{n-1} + E_{A;n-1\to n}$. Here n is the number of base pairs in the nucleus. The enthalpy change, ΔH°_{n-1} can be estimated from the values in Table 8-4. The activation energy for propagating the helix an additional base pair is expected to be similar to that for single strand stacking, roughly 5 kcal mol^{-1}. An estimate of the size of the nucleus can be made by adding appropriate parameters until the measured activation energy is approximated. For example, the measured activation energy for duplex formation by A_6U_6 is -3 kcal mol^{-1} (Craig et al., 1971). Formation

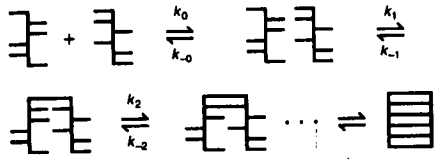

Figure 8-8
Mechanism for duplex formation. Step 0 is alignment of strands without inter-strand bonding. Subsequent steps involve formation of base pairs by hydrogen bonding.

Table 8.8
Kinetic Parameters for Duplex Formation by Oligonucleotides

Sequence	Reference	[Na^+] (M)	t °C	k_{on} ($10^6 M^{-1}s^{-1}$)	k_{off} (s^{-1})	$E_{A,on}$ (kcal mol^{-1})	$E_{A,off}$ (kcal mol^{-1})
$A_6 + U_6$	Pörschke and Eigen (1971).	0.05	23.3	0.53	2300	−8	36
$A_{11} + U_{11}$	Pörschke and Eigen (1971).	0.05	23.3	0.50	210	−12	53
$A_{14} + U_{14}$	Pörschke and Eigen (1971).	0.05	23.5	0.61		−8	
A_3U_4	Craig et al. (1971).	0.25	21	1	3000	−6	37
A_3U_5	Craig et al. (1971).	0.25	21	2	2000	−4	50
A_6U_6	Craig et al. (1971).	0.25	21	1.5	15	−3	60
A_7U_7	Craig et al. (1971).	0.25	21	0.8	1	+5	65
	Breslauer and Bina–Stein (1977).	1	22.1	2.7	3.0	−6	45
A_2GCU_2	Pörschke et al. (1973).	0.05	23.3	1.6	450	3	33
	Pörschke et al. (1973).	1	23	10	200	8	33
A_3GCU_3	Pörschke et al. (1973).	0.05	23.3	0.75	3	7	50
A_4GCU_4	Pörschke et al. (1973).	0.05	23.3	0.13	1.5	8	26
	Pörschke et al. (1973).	1	23	0.9	1	9	26
$A_4G_2 + C_2U_4$	Pörschke et al. (1973).	0.05	16.8	11.4	320	13	34
$A_3G_2 + C_2U_3$	Pörschke et al. (1973).	0.05	23.3	4.4	340	7	43
$CA_3G + CU_3G$	Nelson and Tinoco (1982).	1	21.1	4.6	330	0	39
$dCA_3G + dCT_3G$	Nelson and Tinoco (1982).	1	20	9	70	−0.5	43
GGGC + GCCC	Podder (1971).	0.1	21.5	5.4	40	4.6	29
dGCGCGC	Freier et al. (1983a).	1	25	12.0	0.31	0.8	57
$A_9 + dT_9$	Hogrefe and Maass (1971).	1	23	10	240	−2	43
dCGTGAATTCGCG	Chu and Tinoco (1983).	0.17	31.8	0.08	1.0	22	68
dCGCAGAATTCGCG	Chu and Tinoco (1983).	0.17	31.8	0.07	0.7	16	74
dGCATGC	Williams et al. (1989).	0.042	31.1	0.98	3.9	15	47
		0.042	31.1	1.6	2.1		
		1	31.1	9.9	2.2	−5	40
dGCATGC(0.01 M Mg^{2+})	Williams et al. (1989).	0.012	31.1	7.3	2.4	−3	43

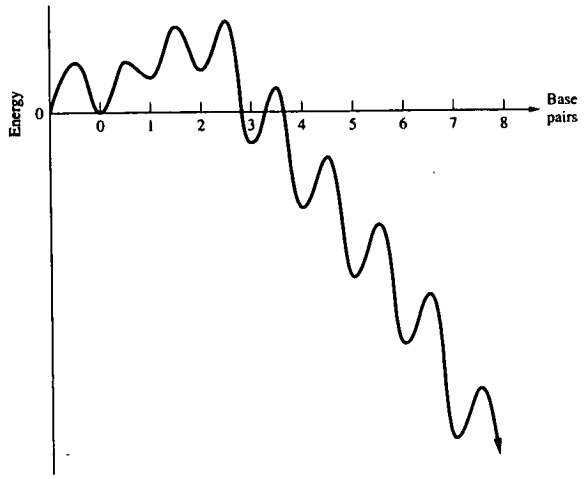

Figure 8-9
Energy profile corresponding to mechanism in Figure 8-9 for formation of duplex from separated single strands.

of four AU base pairs is associated with a $\Delta H^\circ \approx 3.6 + 2(3.7) + 3(-6.8) = -9.4$ kcal mol^{-1}. If addition of the next base pair requires an activation energy of 5 kcal mol^{-1}, then the overall activation energy would be $E_{A,on} = -9.4 + 5 = -4.4$ kcal mol^{-1}. Thus the results are consistent with a nucleus of 5 bp. Since several assumptions are required, the estimate for the size of the nucleus is rough. In general, the results in Table 8-8 are consistent with nuclei of 5 ± 1 bp for oligomers containing only AU pairs, and 2 ± 1 bp for oligomers containing at least two GC pairs.

The data in Table 8-8 can also be used to estimate a rate constant of about 10^7 s^{-1} for addition of the next base pair to the nucleus (Pörschke and Eigen, 1971; Craig et al., 1971). This rate constant is similar to that for single-strand stacking, which supports the choice of activation energy used above.

The activation energies reported for homo-duplex formation by dCGTGAATTCGCG and dCGCAGAATTCGCG are larger than expected from the above analysis (Chu and Tinoco, 1983). Both these oligomers are long, self-complementary, and contain non-Watson–Crick pairings. One possibility is that each oligomer has intramolecular base pairs that must be broken before duplexes can form. Another is that non-Watson–Crick pairings may introduce effects that are not easily predicted from available information. Further experiments are required to sort out these effects.

The reverse rates in Table 8-8 cover a wide range and have large activation energies. These results are consistent with the model described above. Dissociation of duplex requires breaking enough base pairs to return to the nucleus. The energy required for this will depend on sequence and length. If the size of the nucleus and the activation energy for losing a base pair from the nucleus are known, then the ΔH°values in Table 8-4 can be used to predict $E_{A,off}$. For example, for A_6U_6, 7 bp must be broken to return to the 5 bp nucleus. The activation energy for the reverse rate is predicted to be $E_{A,off} = 6 \times 6.8 + 9.4 + (5 + 6.8) = 62$ kcal mol^{-1}. The measured value is 60 kcal mol^{-1} (Craig et al., 1971).

2.3.2 *Predicting Kinetics of Duplex Formation*

The results described above suggest a way to roughly predict the kinetics of duplex formation between oligonucleotides that do not have intramolecular base pairing. At high Na^+ concentration or in the presence of Mg^{2+}, the forward rate is almost always $10^6 M^{-1}s^{-1}$ within an order of magnitude, and the temperature dependence is modest. The equilibrium constant for duplex formation can be predicted from the parameters in Table 8-4, since $K = \exp(-\Delta G^\circ/RT)$ and $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$. The predicted values of K and k_{on} can then be used to predict k_{off} since $K = k_{on}/k_{off}$. For example, the predicted equilibrium constants at 37°C for formation of $(GGCC)_2$ and $(GGCCGGCC)_2$ are 6.6×10^3 and 3.1×10^{12}, respectively, giving predicted off rate constants of 1.5×10^2 and 3×10^{-7} s^{-1}, respectively. The half lives for dissociation of these helixes are predicted from $t_{1/2} = \ln(2)/k_{off} = 0.693/k_{off}$ to be about 5 ms and 27 days, respectively, at 37°C. Note that the off rate is very temperature dependent. For example, at 0°C, the half lives for $(GGCC)_2$ and $(GGCCGGCC)_2$ are predicted to be about 20 s and 30 million years. This type of rough approximation is often sufficient for designing experiments.

3. DOUBLE HELIX FORMATION BY OLIGONUCLEOTIDES WITH LOOPS

Loops are regions of non Watson-Crick paired nucleotides flanked by one or more double helixes. The known varieties of loops are illustrated in Figure 8-11. All occur in structures of RNA (e.g., see Fig. 8-1). While natural DNA is primarily double helix, loops can occur transiently or even be favored under certain conditions. Thus it is important to understand the properties of loops.

There are several experimental ways to determine the thermodynamic properties of loops. All involve measuring properties for structures containing the loop of interest and subtracting out other contributions. For example, ΔG° for the bulge loop in the duplex formed by GCGAGCG + CGCCGC is obtained by taking ΔG° for this bulge duplex and subtracting ΔG° for the fully paired duplex formed by GCGGCG + CGCCGC:

$$\Delta G^\circ_{bulge} = \Delta G^\circ(\text{GCGAGCG} + \text{CGCCGC}) - \Delta G^\circ(\text{GCGGCG} + \text{CGCCGC})$$

The assumption is that the loop does not affect the stabilities of other regions so the free energies are additive. This assumption may not be very good (Longfellow et al., 1990). Often loops are studied in structures formed by single strands, for example the hairpin formed by $A_6C_6U_6$. If comparisons in such cases are made with helixes formed by bimolecular associations, it is necessary to correct for initiation of the bimolecular helix first. Unfortunately, fewer data are available for loops than for fully base paired helixes. Much of the available data is discussed below. Examples of calculations of thermodynamic properties for structures containing loops can be found in Xia et al. (1999).

3.1 Bulge Loops

Bulge loops are loops in which unpaired nucleotides occur on only one strand of a double helix (Fig. 8-11). In structures of natural RNA, the most common bulge contains one nucleotide, and bulges are known to be important for protein binding and tertiary folding (Peattie et al., 1981; Romaniuk et al., 1987; Flor et al., 1989; Cate et al., 1996, 1997). In DNA, bulges may be important in frameshift mutagenesis in sequences with repeating base pairs (Okada et al., 1972), and detection of bulges by gel electrophoresis can reveal mutations (Triggs-Raine and Gravel, 1990; Rommens et al., 1990). Chemical modification experiments on large RNA

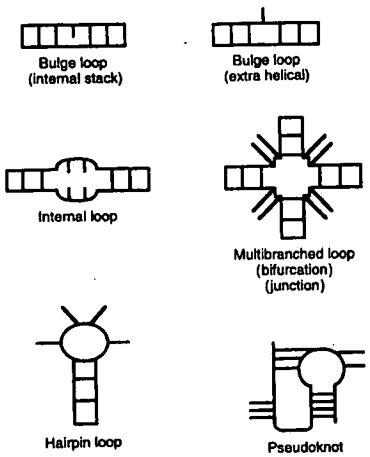

Figure 8-10
Schematic of various types of loops.

molecules (Moazed et al., 1986) suggest the structure of a bulge may be sequence dependent. Using R and Y to denote purines and pyrimidines, respectively, bulged nucleotides in N^RR sequences are usually moderately or strongly reactive. Bulged nucleotides in Y^RY sequences are usually protected from modification. NMR experiments indicate the bulged As in $(dCGC\underline{A}GAATTCGCG)_2$, $(dCGC\underline{A}GAGCTCGCG)_2$ and $(dCGCGAAATTT\underline{A}CGCG)_2$ are intercalated in the helix (Patel et al., 1982, 1986; Hare et al., 1986), while the bulged C in d(CAAA$\underline{C}$AAAG)·d(CTTTTTTG) and the bulged U in r(CUGG$\underline{U}$GCGG)·r(CCGCCCAG) are extrahelical (Morden et al., 1983; van den Hoogen et al., 1988) (see Chapter 4). Thermodynamic parameters for single nucleotide bulges indicate they destabilize the helix, but with little dependence on the identity of the bulged base (Longfellow et al., 1990; Groebe and Uhlenbeck, 1989). The destabilization, however, does depend on more than the adjacent base pairs (Longfellow et al., 1990). Thus the nearest neighbor model is oversimplified for bulges. A limited number of sequences have been studied, and representative data are listed in Table 8-9.

Some bulges occur in sequences that allow migration of the unpaired nucleotide. For example, any of the middle C nucleotides in the lower strand of $^{dGATGGG\text{-}CAG}_{CTACCCCGTC}$ could be bulged. The NMR experiments indicate such migration occurs at a rate of at least 100–1000 s^{-1} (Woodson and Crothers, 1987).

Theoretical considerations indicate bulge loops should be more destabilizing as the number of unpaired nucleotides increases. While only a few measurements have been made on bulges larger than 1, the results are mostly in agreement with this prediction (see Table 8-9) (Longfellow et al., 1990; Yuan et al., 1979; Weeks and Crothers, 1993).

3.2 Internal Loops

An internal loop is formed when a double helix is interrupted by non-Watson–Crick paired nucleotides on both strands (Fig. 8-11). The size of an internal loop is the total number of nucleotides in the loop. Internal loops can be symmetrical or asymmetrical with respect to the number of loop residues on each strand. Asymmetric internal loops are less stable than symmetric internal loops (Peritz et al., 1991). A particularly important subclass of internal loop, the mismatch, contains two nucleotides. Stabilities of mismatches are important for determining stabilities of duplexes formed between DNA hybridization probes and target sequences that are not completely complementary. Due to redundancy in the genetic code, this situation often occurs when the sequence of a probe is designed from the sequence of a gene's protein product. Modified bases, like inosine, that are less discriminating than A, C, G, and T, can be used in such cases to lessen the effect of the mismatch (Martin et al., 1985; Cheong et al., 1988; Nichols et al., 1994). Formation of large internal loops is the first step in melting of DNA (Blake and Decourt, 1987; Wada et al., 1980; Gotoh, 1983; Wartell and Benight, 1985). In RNA, internal loops of many sizes are known to occur (see Fig. 8-1). These loops allow structural bends (Murphy and Cech, 1993) and tertiary interactions (Costa and Michel, 1995), and can also form binding sites (Sassanfar and Szostak, 1993; Fan et al., 1996; Jiang et al., 1996; Dieckmann et al., 1996; Yang et al., 1996).

Considering the wide variety of possible internal loops, relatively few experimental data are available. Some measurements involve single mismatches in DNA

Table 8.9
Free Energy Increments at 37°C for Bulge Loops[a]

Bulge Sequence	Reference	[Na^+] (M)	t_m(°C $10^{-4}M$)	Δt_m(°C $10^{-4}M$)	$\Delta\Delta G^\circ_{bulge,37}$ (kcal mol^{-1})
RNA					
poly(A,A°)·polyU	Fink and Crothers (1972)	0.2–0.5			2.9
5'GCG^{A}GCG3' 3'CGC-CGC5'	Longfellow et al. (1990)	1	39	20	3.5
5'GCG^{U}GCG3' 3'CGC-CGC5'	Longfellow et al. (1990)	1	38	21	3.7
5'GCG-GCG3' 3'CGC$_{A}$CGC5'	Longfellow et al. (1990)	1	39	20	3.5
5' GCG^{A}GCGA3' 3'ACGC-CGC5'	Longfellow et al. (1990)	1	48	22	5.1
5' GCG^{A}GUCA3' 3'ACGC-CAG5'	Longfellow et al. (1990)	1	49	16	2.0
5'^{G}GGACUC^{A}CG^{AU} 3'$_{UA}$UCUGAG-GC$_{A}^{U}$	Groebe and Uhlenbeck (1989)	1	79	16	3.4
5'^{G}GGACUC^{O}CG^{AU} 3'$_{UA}$UCUGAG-GC$_{A}^{U}$	Groebe and Uhlenbeck (1989)	1	79	16	3.5
5'^{G}GGACUC^{U}CG^{AU} '$_{UA}$UCUGAG-GC$_{A}^{U}$	Groebe and Uhlenbeck (1989)	1	80	14	2.7
5'GCG^{AA}GCG3' 3'CGC—CGC5'	Longfellow et al. (1990)	1	24	35	5.2
5'GCG^{AAA}GCG3' 3'CGC—CGC5'	Longfellow et al. (1990)		14	45	5.7
5'GCG—GCG3' 3'CGC$_{AAA}$CGC5'	Longfellow et al. (1990)	1	38	21	3.7
DNA					
5'dCA$^{C}_{1}$A$_1$G3' 3' GT$_1$-T$_1$C5'	Morden et al. (1983)	1	18	14	2.6
5'dCTG^{A}CCCATC3' 3' GAC-GGGTAG5'	Woodson and Crothers (1987)	0.1	31	15	3.1
5'dCTG^{C}CCCATC3' 3' GAC GGGTAG5'	Woodson and Crothers (1987)	0.1	37	10	2.1

[a]The parameters Δt_m and $\Delta\Delta G^\circ_{bulge,37}$ are differences between the duplexes with and without the bulge.

oligonucleotides (Gaffney and Jones, 1989; Martin et al., 1985; Tibanyenda et al., 1984; Aboul-ela et al., 1985; Allawi and SantaLucia, 1997, 1998). Table 8-10 lists some of these results. Substituting a mismatch for a base pair always gives a less stable duplex. In general, the destabilization is associated with a less favorable ΔH° and more favorable ΔS° of duplex formation, which is consistent with the expectation of reduced bonding in a mismatch coupled with increased flexibility.

Data for RNA internal loops indicate that internal loops become more destabilizing as they increase in size (Weeks and Crothers, 1993; Peritz et al., 1991; Gralla and Crothers, 1973). For example, the melting temperatures at $10^{-4}M$ strand for

Table 8.10
Thermodynamic Parameters for Duplex Formation in 1 M NaCl by DNA Oligonucleotides Containing Mismatches[a]

Sequence	Mismatch	$-\Delta H^\circ$ (kcal mol^{-1})	$-\Delta S^\circ$ (eu)	$-\Delta G^\circ_{37}$ (kcal mol^{-1})	t_m(°C at $10^{-4}M$)
$CA_1CA_1G + CT_1GT_1G$		64.5	183	7.7	42.9
$CA_1GA_1G + CT_1CT_1G$		62.8	179	7.3	40.8
$CA_1AA_1G + CT_1TT_1G$		68.0	196	7.2	40.1
$CA_1TA_1G + CT_1AT_1G$		58.6	168	6.5	36.8
$A_1GA_1G + CT_1GT_1G$	GG	53.5	158	4.5	25.6
$CA_1TA_1G + CT_1GT_1G$	TG	55.6	165	4.4	25.7
$CA_1GA_1G + CT_1AT_1G$	GA	52.6	156	4.2	23.9
$CA_1GA_1G + CT_1TT_1G$	GT	46.7	137	4.2	22.3
$CA_1AA_1G + CT_1GT_1G$	AG	39.9	116	3.9	18.0
$CA_1AA_1G + CT_1AT_1G$	AA	36.9	107	3.7	15.0
$CA_1CA_1G + CT_1TT_1G$	CT	53.2	161	3.3	19.1
$CA_1TA_1G + CT_1CT_1G$	TC	50.0	151	3.2	17.5
$CA_1CA_1G + CT_1AT_1G$	CA	(40.3)	(120)	(3.1)	(13)
$CA_1TA_1G + CT_1TT_1G$	TT	(54.6)	(167)	(2.8)	(17)
$CA_1AA_1G + CT_1CT_1G$	AC	(35.8)	(106)	(2.9)	(9)
$CA_1CA_1G + CT_1CT_1G$	CC	(55.3)	(171)	(2.3)	(15)
5'CAACTTGATATTAATA					
3'GTTGAACTATAATTAT		102.1	289	12.4	55.8
5'CAACTTGATATTAATA					
3'GTTGAGCTATAATTAT	TG	92.6	266	10.1	49.4
5'CAACTTGATATTAATA					
3'GTTGAACTATAGTTAT	TG	95.5	274	10.5	50.5
5'CAACTTGATATTAATA					
3'GTTGAACTCTAATTAT	TC	98.4	286	9.7	47.3
5'CAACTTGATATTAATA					
3'GTTGAATTATAATTAT	GT	91.3	264	9.4	47.1
5'CAACTTGATATTAATA					
3'GTTGAACCATAATTAT	AC	90.9	265	8.7	44.6
5'CAACTTGATATTAATA					
3'GTTGAACAATAATTAT	AA	92.0	267	9.26	46.2

[a]Data for $CA_1XA_1G + CT_1YT_1G$ sequences are from Aboul-ela et al. (1985). Data in parentheses are significantly less accurate. Other sequences are from Tibanyenda et al. (1984). Similar results for d(GGTTXTTGG) + d(CCAAYAACC) have been reported by Gaffney and Jones (1989).

$(CGCA_nGCG)_2$ with $n = 0$, 1, 2, and 3, are 58, 40, 36, and 32°C, respectively (Peritz et al., 1991). The stability of an internal loop is also dependent on the sequence in the loop (SantaLucia et al., 1991; Wu et al., 1995; Schroeder et al., 1996; Xia et al., 1997). For example, Table 8-11 lists ΔG° values for internal loops containing adjacent, identical mismatches. For this case, GU, GA, and UU double mismatches often stabilize a duplex, while other double mismatches destabilize a duplex. This sequence

Table 8.11
Free Energy Increments (ΔG°_{37} in kcal mol^{-1}) for Tandem Mismatches in RNA Oligonucleotides in 1 *M* NaCl[a]

Mismatches → / ←	UG / GU	GU / UG	GA / AG	AG / GA	UU / UU	GG / GG	CA / AC	CU / UC	UC / CU	CC / CC	AC / CA	AA / AA
Closing bp												
5'G 3'C	−4.9	−4.1	−2.6	−1.3	−0.5		1.0	1.1			0.9	1.5
5'C 3'G	−4.2	−1.1	−0.7	−0.7	−0.4	0.8	1.1	1.4	1.4	1.7	2.0	1.3
5'U 3'A	−2.6	−0.3	0.7		1.1		[1.9][c]	2.2	2.8			2.8
5'A 3'U	−1.9	0.2	0.3		0.6		2.3				2.5	2.8

[a]Free energy increments are calculated as in the following example:

$$\Delta G^\circ_{37,\text{loop}}\begin{pmatrix}\text{CAGG}\\ \bullet\ \ \bullet\\ \text{GGAC}\end{pmatrix} = \Delta G^\circ_{37}(\text{CGC}\underline{\text{AG}}\text{GCG}) - \Delta G^\circ(\text{CGCGCG}) + \Delta G^\circ_{37}\begin{pmatrix}\text{CG}\\ \text{GC}\end{pmatrix}$$

Most sequences had 3 bp on each side of the internal loop. Some values are averages from more than one sequence.

[b]He et al., (1991); SantaLucia et al., (1991); Wu et al., (1995); Mathews et al., (1999).

[c]Sequence with this tandem mismatch had either an unusual conformation or mixture of conformations.

dependence is thought to be due to hydrogen bonding in the stable loops (SantaLucia et al., 1991; Wu et al., 1995). The NMR studies of internal loops have revealed a variety of noncanonical interactions indicating that rules for thermodynamic stability may be complex (Wimberly et al., 1993; SantaLucia and Turner, 1993; Peterson et al., 1994; Battiste et al., 1994; Wu and Turner, 1996; Wu et al., 1997).

The most common mismatch in RNA is GU, and often GU mismatches are conserved (Gutell et al., 1994; Michel and Westhof, 1990). This conservation may imply importance in tertiary and functional interactions, and this has been shown for group I introns (Pyle et al., 1994; Knitt et al., 1994; Strobel and Cech, 1995). Table 8-12 lists thermodynamic parameters for GU mismatches (Mathews et al., 1999). Note that the nearest neighbor, 5'GU3'/5'GU3', is destabilizing in most but not all contexts. This sequence dependence is a non-nearest neighbor effect. The destabilizing motif rarely occurs in secondary structures of natural RNA. Additional terms are applied for terminal GU pairs, and these terms are assumed to be the same as those listed for terminal AU pairs in Table 8.4.

Table 8.12
Thermodynamic Parameters for Helix Propagation by G:U Pairs in RNA Oligonucleotides in 1 *M* NaCl[a]

Propagation Sequence → / ←	ΔH° (kcal mol^{-1})	ΔS° (eu)	ΔG°_{37} (kcal mol^{-1})
GC / UG	−12.59	−32.5	−2.51
GG / UC	−12.11	−32.2	−2.11
GG / CU	−8.33	−21.9	−1.53
GA / UU	−12.83	−37.3	−1.27
GU / UA	−8.81	−24.0	−1.36
CG / GU	−5.61	−13.5	−1.41
UG / AU	−6.99	−19.3	−1.00
AG / UU	−3.21	−8.6	−0.55
UG / GU	−9.26	−30.8	+0.30
GU / UG	−14.59(−14.14)[b]	−51.2(−42.2)[b]	+1.29(−1.06)[b]
GG / UU	−13.47	−44.9	+0.47

[a]He et al., (1991); Mathews et al. (1999). Additional terms are applied for terminal GU pairs, and these are assumed to be the same as those listed for terminal AU pairs in Table 8.4.

[b]$\begin{matrix}\text{GU}\\ \text{UG}\end{matrix}$ in the contexts $\begin{matrix}\text{AGUU}\\ \text{UUGA}\end{matrix}$, $\begin{matrix}\text{CGUG}\\ \text{GUGC}\end{matrix}$, and $\begin{matrix}\text{UGUA}\\ \text{AUGU}\end{matrix}$ has a ΔG°, ΔH°, and ΔS° of +1.29 kcal mol^{-1}, −14.59 kcal mol^{-1}, and −51.2 eu, respectively, but in the context $\begin{matrix}\text{GGUC}\\ \text{CUGG}\end{matrix}$ it has a ΔG°, ΔH°, and ΔS° of −1.06 kcal mol^{-1}, −14.14 kcal mol^{-1}, and −42.2 eu, respectively.

3.3 Hairpin Loops

Hairpin loops occur when nucleic acid strands fold back on themselves to make base pairs (Fig. 8-11). Hairpins are widespread in structures of natural RNA (see Fig. 8-1). In DNA, there are many instances of sequences capable of forming hairpins as alternates to the fully base paired double helix. There is evidence that such hairpins will form when the DNA is placed under superhelical tension (Gellert et al., 1979; Panayotatos and Wells, 1981; Sullivan and Lilley, 1986).

Conformational changes involving hairpins are intramolecular. Thus thermodynamic parameters can be determined from calorimetric and spectroscopic data in a manner similar to that described for single-strand stacking in Section 8.1.1. Fortunately, hairpin transitions are more cooperative than those of single strands, thus simplifying determination of baselines. For transitions considered two-state, $\Delta H°$ is simply derived from (Bloomfield et al., 1974):

$$\text{Unimolecular} \qquad \Delta H° = 4RT_m^2 \left(\frac{\partial \alpha}{\partial T}\right)_{T=T_m} \tag{8-22a}$$

Here the partial derivative is the slope at the T_m of a plot of fraction of strands in single-strand state versus temperature. The value of $\Delta S°$ can then be determined from Eq. 8-7b. Equation 8-22a is a specific case of the general equation for an equilibrium with N strands (Marky and Breslauer, 1987):

$$N\text{-mer} \qquad \Delta H° = (2 + 2N)RT_m^2 \left(\frac{\partial \alpha}{\partial T}\right)_{T=T_m} \tag{8-22b}$$

More data are available for hairpins than for other loops (Table 8-13; Hilbers et al., 1994). The free energy increment for loop formation is unfavorable, and the minimum loop size is 2 (Orbons et al., 1986, 1987; Blommers et al., 1989; Jucker and Pardi, 1995). The stability of hairpins is dependent on loop size, the sequence in the loop, and the base pair closing the loop (see Table 8-13). For example, the CG closed tetraloop formed by CUUCGG is unusually stable (Tuerk et al., 1988; Antao and Tinoco, 1992). The complete basis for the sequence dependence of hairpin stability has not yet been unraveled (Hilbers et al., 1994), although hydrogen bonding and stacking within the loop probably make contributions (Varani et al., 1991; SantaLucia et al., 1992; Serra et al., 1994). An equation that provides a reasonable prediction for the free energy increment in kilocalories per mole associated with RNA hairpin loops greater than or equal to 4 nucleotides is (Serra et al., 1994, 1997; Mathews et al., 1999):

$$\Delta G°_{37,\text{loop}} = \Delta G°_{37,\text{length}} + \Delta G°_{37,\text{mm}} - 0.8 \text{ if first mismatch is GA or UU} \tag{8-23}$$

Here $\Delta G°_{37,\text{length}}$ is the length dependence of hairpin $\Delta G°$ (see Table 8-14) and $\Delta G°_{37,\text{mm}}$ is the free energy increment for the first mismatch in the loop (Table 8-6). The sarcin/ricin loop from 28S ribosomal RNA has 17 nucleotides, many of which are involved in noncanonical interactions (Szewczak et al., 1993). This result suggests that the complete set of rules for hairpin stability may be complex.

Table 8.13
Thermodynamic Parameters for Hairpin Loops

Stem Sequence	Loop Sequence	Reference	Salt	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ eu	$\Delta G°_{37}$ (kcal mol^{-1})	t_m °C
RNA							
GGGAUAC • • • • • ACCUAUG	A3	Groebe and Uhlenbeck (1988).	1M Na^+	−56.2	160.9	−6.3	76.4
	A4	Groebe and Uhlenbeck (1988).	1M Na^+	−63.8	−181.2	−7.6	78.9
	A5	Groebe and Uhlenbeck (1988).	1M Na^+	−64.6	−183.4	−7.7	79.1
	A7	Groebe and Uhlenbeck (1988).	1M Na^+	−61.5	−175.1	−7.2	78.1
	A9	Groebe and Uhlenbeck (1988).	1M Na^+	−52.6	−151.7	−5.6	73.3
	U3	Groebe and Uhlenbeck (1988).	1M Na^+	−60.1	−171.1	−7.0	78.1
	U4	Groebe and Uhlenbeck (1988).	1M Na^+	−67.4	−191.0	−8.2	79.7
	U5	Groebe and Uhlenbeck (1988).	1M Na^+	−67.4	−191.1	−8.1	79.5
	U7	Groebe and Uhlenbeck (1988).	1M Na^+	−66.0	−189.0	−7.4	76.1
	U9	Groebe and Uhlenbeck (1988).	1M Na^+	−73.8	−212.0	−8.0	74.3
GGU • • • CCA	AUAAUA	Serra et al. (1993).	1M Na^+	−25.9	−82.4	−0.3	40.7
GGA • • • CCU	AUAAUA	Serra et al. (1993).	1M Na^+	−29.2	−91.8	−0.7	44.2
GGG • • • CCC	AUAAUA	Serra et al. (1993).	1M Na^+	−33.5	−101.0	−2.2	58.7
GGC • • • CCG	AUAAUA	Serra et al. (1993).	1M Na^+	−34.3	−101.7	−2.7	64.0
GGC • • • CCG	GUAAUA	Serra et al. (1993).	1M Na^+	−36.9	−107.8	−3.4	68.5
GGAC • • • • CCUG	UUCG	Antao and Tinoco (1992).	1M Na^+	−55.9	−159.9	−6.3	76.2
GGAC • • • • CCUG	UUUG	Antao and Tinoco (1992).	1M Na^+	−44.0	−128.0	−4.3	70.3
GGAG • • • • CCUC	UUCG	Antao and Tinoco (1992).	1M Na^+	−44.8	−131.4	−4.0	67.7
CA GA • • • • $GUm^5\psi$	$CmUG_m$ AYA^A	Clore et al. (1994).	0.5M K^+	−38	−118	−1.4	53
DNA							
GGAC • • • • CCTG	TTCG	Antao and Tinoco (1992).	1M Na^+	−31.3	−93.9	−2.2	60.4
GGAC • • • • CCTG	TTTG	Antao and Tinoco (1992).	1M Na^+	−31.0	−93.0	−2.2	59.8
GGAG • • • • CCTC	TTCG	Antao and Tinoco (1992).	1M Na^+	−30.2	−92.5	−1.5	53.8
ATCCTA • • • • • • TAGGAT	T_4	Hilbers et al. (1985).	0.2M Na^+	−39	−119.2		54
	T_5	Hilbers et al. (1985).	0.2M Na^+	−43	−132.7		51
	T_6	Hilbers et al. (1985).	0.2M Na^+	−43	−133.9		48
	T_7	Hilbers et al. (1985).	0.2M Na^+	−43	136.0		44
CGAACG • • • • • • GCTTGC	A_4	Senior et al. (1988a).	0.1M Na^+	−33.7			63.4
	G_4	Senior et al. (1988a).	0.1M Na^+	−41.8			65.6
	C_4	Senior et al. (1988a).	0.1M Na^+	−46.3			65.8
	T_4	Senior et al. (1988a).	0.1M Na^+	−49.3			67.4
m^5C Gm^5C • • • Gm^5C G	GA	Orbons et al. (1987).		−28.9	87.8		56

Table 8.14
Free Energy Increments (kcal mol^{-1}) for RNA Loops at 37°C in 1 *M* NaCl[a]

Loop Size	Internal Loop	Bulge Loop	Hairpin Loop
1		+3.8	
2	+0.4[b]	+2.8	–
3	[c]	+3.2	+5.7
4	[c]	(+3.6)	+5.6
5	+1.8	(+4.0)	+5.6
6	+2.0	(+4.4)	+5.4
7			+5.9
8			+5.6
9			+6.4

[a]From Mathews et al. 1999. For larger loop sizes, n, use $\Delta G^\circ(n) = \Delta G^\circ(n_{max}) + 1.75RT \ln(n/n_{max})$, where n_{max} is 6, 6, and 9 for internal, bulge, and hairpin loops, respectively. Parameters not derived from experimental measurements are listed in parentheses. Parameters for hairpin loops are for closure by CG or GC, and for loops greater than or equal to 4 assume additional stability is conferred by terminal mismatches at helix ends (see Table 8-6). A reasonable approximation for hairpin loops greater than or equal to 4 is given by: $\Delta G^\circ_{37,loop} = \Delta G^\circ_{37,length} + \Delta G^\circ_{37,mm} - 0.8$ (if first mismatch is UU or GA (not AG)). Asymmetric internal loops with branches of N1 and N2 nucleotides must be penalized additionally by the minimum of 3.0 or 0.5 |N1 – N2| kcal mol^{-1}. For internal loops larger than 4, each terminal GA and AG mismatch is given a favorable free energy increment of –1.1 kcal mol^{-1} and each terminal UU is given a bonus of –0.7 kcal mol^{-1}. For each AU or GU pair closing an internal loop, a penalty of 0.2 kcal mol^{-1} is applied in addition to the 0.45 kcal mol^{-1} penalty for terminating a helix with AU or GU. The parameter for bulge loops of one nucleotide is based on the assumption that additional stability is conferred by stacking of the adjacent base pairs as approximated by nearest neighbor parameters (see Table 8-4). It is assumed there is no stacking across bulges of two or more nucleotides.

[b]Rough average for single mismatches that are not GG and that have 2 adjacent GC pairs. The equivalent value for GG mismatches is –2 kcal mol^{-1}.

[c]Free energies of internal loops with 3 or 4 nucleotides are estimated with special rules (Xia et al., 1999; Mathews et al., 1999).

The kinetics of hairpin formation have been studied by temperature-jump methods (Pörschke, 1974; Riesner et al., 1973; Coutts, 1971; Coutts et al., 1974; Orbons et al., 1986, 1987). Forward rates are between 10^4 and 10^5 s^{-1} and change little with temperature. As with duplex formation, this result allows rough prediction of the rate of unfolding for a hairpin, if the equilibrium constant for hairpin formation is known or can be predicted. For example, at the melting temperature, $K = k_{for}/k_{rev} = 1$. Thus the rate of unfolding of a hairpin at the T_m is also roughly $10^4 - 10^5$ s^{-1}. In a temperature-jump experiment, the corresponding relaxation time is given by $\tau^{-1} = k_{for} + k_{rev}$, or about 5–50 μs at the T_m.

3.4 Multibranch Loops (Junctions)

Multibranch loops (junctions) occur when three or more helixes intersect (Fig. 8-11). The term loop is somewhat misleading since it is possible for these structures to exist with no unpaired nucleotides. Models for such a structure with four helixes have been studied with DNA oligonucleotides (Marky et al., 1987; Lu et al., 1992). The ΔH° for formation of one structure was simply the sum of the ΔH° values for formation of each separate helix in the structure. Thus the junction did not significantly perturb the bonding in attached helixes (Marky et al., 1987). The thermodynamic parameters for forming a four-arm junction structure from two paired duplexes are $\Delta G^\circ_{18} = 1.1$ kcal mol^{-1} and $\Delta H^\circ = 27.1$ kcal mol^{-1} (Lu et al., 1992). A DNA junction with three helixes is stabilized by adding unpaired nucleotides to the junction (Leontis et al., 1991). Melting of tRNA involves changes in the size of a multibranch loop (Riesner and Römer, 1973; Crothers and Cole, 1978; Crothers et al., 1974).

3.5 Knotted Loops

If two nucleotides i and j are base paired, a pseudo knot is formed when a nucleotide between i and j in the primary sequence is base paired with a nucleotide not between i and j (Fig. 8-11). These structures were first postulated to explain similar biochemical reactivities for tRNA and the 3′ ends of certain viral RNAs from plants (Pleij et al., 1985). They have subsequently been shown to occur in oligonucleotides of appropriate sequence (Wyatt et al., 1990). It was suggested that the minimum loop sizes for the 5′ and 3′ loops are two and three nucleotides, respectively, for A-form RNA, and perhaps even one or two for a distorted structure (Pleij et al., 1985). Studies of oligonucleotide models are consistent with the predicted minimum loop sizes for A form, and indicate pseudoknot stability is dependent on loop size and sequence (Wyatt et al., 1990). The presence of Mg^{2+} or high Na^+ concentrations also stabilizes pseudoknots (Wyatt et al., 1990).

3.6 Large Loops

Loops larger than 12 nucleotides are rarely seen in nucleic acid structures. Nevertheless, the properties of large loops are important for successful predictions of nucleic acid structure, and for predicting the melting behavior of nucleic acids. Unfortunately, it is difficult to experimentally determine the properties of large loops, especially given the number of possible permutations. Theoretical considerations, however, provide a reasonable approximation for the ΔG° for initiation of large loops (Jacobson and Stockmayer, 1950). Considerations of conformational entropy effects lead to the following equation for the ΔG° of initiation of a large loop with N nucleotides:

$$\Delta G^\circ_{loop}(N) = \Delta G^\circ_{loop}(n) + aRT \ln(N/n) \qquad (8\text{-}24)$$

Here $\Delta G^\circ_{loop}(n)$ is the experimentally determined ΔG° for a loop of n nucleotides, and a is a constant. Theoretical considerations suggest the value of a is about 1.75 (Fisher, 1966). Experiments on bulge and internal loops are consistent with Eq. 8-24

(Weeks and Crothers, 1993). A set of loop parameters derived from experimental and theoretical results is given in Table 8-14.

Reactions involving covalently closed circular nucleic acids provide another situation where the free energies of large loops are probably important. For example, the melting temperature of a covalently closed circular DNA of 26 nucleotides that forms a dumbbell-shaped structure is more than 30°C higher than for the same sequence with a single break in the sugar–phosphate backbone (Erie et al., 1989). This difference corresponds to a more favorable ΔG°_{37} of 7 kcal mol^{-1} for folding into the dumbbell. A large fraction of this favorable free energy is expected from the conformational constraints placed on a melted loop when it is covalently closed (Jaeger et al., 1990). The same effect can also largely account for the enhanced binding observed between oligomers when one is a covalently closed circle (Prakash and Kool, 1992). While such a conformational effect might be expected to provide a more favorable entropy for folding, a more favorable enthalpy is actually observed (Erie et al., 1989; Prakash and Kool, 1992). This apparent anomaly is seen even in reactions of small molecules, however (Jencks, 1975).

4. DOUBLE HELIX FORMATION WITH POLYNUCLEOTIDES

An understanding of helix formation in oligonucleotides provides a basis for understanding the properties of nucleic acid polymers. It is likely that interactions important in oligomers will be major factors determining polynucleotide properties, but that new interactions will also be important. Tertiary interactions are expected to be weaker energetically than secondary structure interactions, and there is some experimental evidence to support this (Crothers et al., 1974; Banerjee et al., 1993; Jaeger et al., 1993; Laing and Draper, 1994). The most extensively studied cases of melting of nucleic acid polymers are tRNA and DNA. In both cases, melting typically starts with formation of large loops.

4.1 Transfer RNA

The experimentally determined pathway for melting of *Escherichia coli* formylmethionine tRNA in 0.17 *M* Na^+ is shown in Figure 8-12 (Crothers et al., 1974). First, tertiary interactions and the dihydrouridine stem melt, followed by the $T\Psi C$ stem, the anticodon stem, and finally the acceptor stem. The order of melting for other tRNAs may be different, indicating a sequence dependence (Riesner and Römer, 1973).

The kinetics of melting for tRNA are similar to expectations from oligomers (Riesner and Römer, 1973; Crothers et al., 1974). The $T\Psi C$ and anticodon hairpins typically melt with relaxation times of 10–100 μs. Melting of the acceptor stem is associated with a relaxation time of about 1 ms, presumably because it is the final stem to melt and therefore opens a large hairpin loop. Melting of the least stable helix, the dihydrouridine stem, is associated with relaxation times of about 1–10 ms. This time

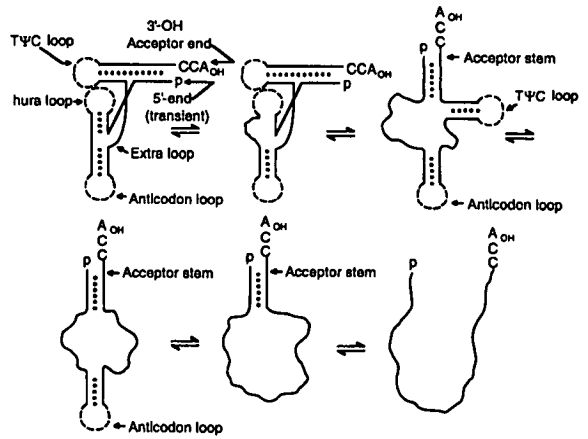

Figure 8-11
Pathway for melting of *E. coli* formylmethionine tRNA in 0.17 *M* Na^+ [Reprinted with permission from Crothers, D. M., Cole, P. E., Hilbers, C. W., and Shulman, R. G., The Molecular Mechanism of Thermal Unfolding of *Escherichia coli* Formylmethionine transfer RNA, *J. Mol. Biol.* **87**, 63–88, Copyright ©1974, by permission of the publisher Academic Press Limited London.] The dihydrouridine helix and tertiary interactions melt first, followed in succession by the $T\Psi C$, anticodon, and acceptor stem helixes.

is much longer than expected from model hairpins, and is associated with concomitant disruption of tertiary structure. This long relaxation time is one example of effects in polymers that are not expected from simple model systems.

4.2 DNA

A typical differentiated melting curve and pathway for melting of DNA are shown in Figure 8-13 (Blake and Decourt, 1987; Wada et al., 1980; Gotoh, 1983; Wartell and Benight, 1985; Lerman et al., 1984). Depending on sequence and length, DNA can begin melting from the ends or the middle. As temperature is raised, AT rich regions melt first, leaving internal loops in long DNA. Typically, loops with nucleotides from 100–350 bp are formed. A single large internal loop of n nucleotides is more favorable than two separate internal loops containing a total of n nucleotides (see Table 8-14). Thus loops coalesce. Occasionally, it is necessary to also consider intermediates with hairpin loops (Wartell and Benight, 1985). The result, as shown in Figure 8-13,

is a series of transitions as the temperature is raised. The parameters in Table 8-4 coupled with parameters for internal and hairpin loops, should be able to predict the melting behavior for any DNA sequence. Alternative sets of parameters have also been suggested (Wada et al., 1980; Gotoh, 1983; Wartell and Benight, 1985; Breslauer et al., 1986; Lerman et al. 1984; Klump, 1990; Delcourt and Blake, 1991; Quartin and Wetmur, 1989; Doktycz et al., 1992). Parameters for loops have been determined by measuring the melting of DNAs that have AT rich regions inserted (Blake and Delcourt, 1987). For sequences longer than about 300 bp, separation of strands can be neglected in these predictions, but it must be included for shorter sequences (Wartell and Benight,

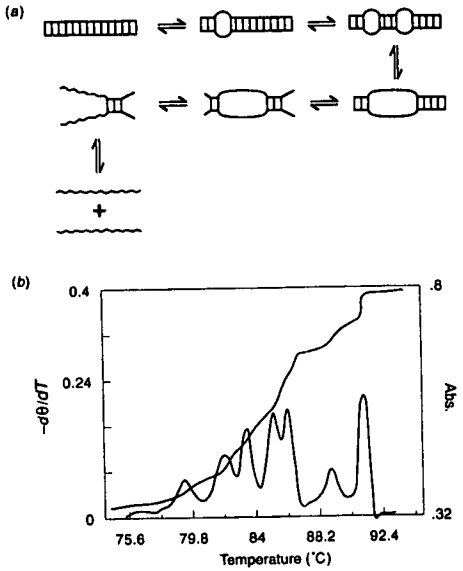

Figure 8-12
(*a*) Typical pathway for melting of DNA. (*b*) Absorbance versus temperature and differentiated melting curves for the 1630 bp Hinf I restriction endonuclease fragment of plasmid pBR322. [Reprinted from Wartell, R. M. and Benight, A. S., Thermal Denaturation of DNA Molecules: A Comparison of Theory with Experiment, *Phys. Rep.*, **126**, 67–107. Copyright ©1985 with permission of Elsevier Science-NL, Amsterdam, The Netherlands.]

1985). In several cases, predictions are reasonably successful (Wartell and Benight, 1985; Steger, 1994).

For long DNA sequences with relatively random distributions of nearest neighbors, simple equations have been developed empirically to approximately predict t_m from the fraction of GC content, F_{GC} (Marmur and Doty, 1962). One that is useful for long DNAs of quasirandom sequence with $0.3 < F_{GC} < 0.7$ and $0.02 \leq [Na^+] \leq 0.4M$ is (Blake, 1996)

$$t_m(°C) = 193.67 - (3.09 - F_{GC})(34.64 - 6.52 \log[Na^+]). \tag{8-25}$$

Another that includes the effect of duplex length, D, and the percentage of mismatches, P, but neglects the effect of GC content on the salt dependence is (Wetmur, 1991)

$$t_m(°C) = 81.5 + 41F_{GC} + 16.6 \log_{10}\left(\frac{[Na^+]}{1.0 + 0.7[Na^+]}\right) - \frac{500}{D} - P \tag{8-26}$$

This equation is valid up to 1 M Na^+. The analogous equation for RNA is

$$t_m(°C) = 78 + 70F_{GC} + 16.6 \log_{10}\left(\frac{[Na^+]}{1.0 + 0.7[Na^+]}\right) - \frac{500}{D} - P \tag{8-27}$$

An interesting application of DNA melting is the use of gels containing denaturant gradients to separate DNAs of similar length but different sequence, and to detect single base pair changes between DNAs (Lerman et al., 1984). These applications arise from a large decrease in gel mobility when denaturation induces an internal loop. Predictions for comparison with experiment are possible since denaturants such as urea and formamide mimic a temperature increase (Lerman et al., 1984; Klump and Burkart, 1977).

The kinetics of association and dissociation for large DNAs is quite different from that of oligomers (Wetmur and Davidson, 1968; Bloomfield et al., 1974; Cantor and Schimmel, 1980; Studier, 1969; Record and Zimm, 1972). This results from sequence complexity and large size. For example, association rates are affected by intramolecular helix formation and by incorrect intermolecular helix formation. Thus, when two separated strands of a large DNA are quickly cooled, they are kinetically trapped in nonnative structures and essentially never able to reform the perfectly matched helix. Under conditions where the perfectly matched helix can be formed, the rate constant k_2 for this association has been found experimentally to be given by (Wetmur and Davidson, 1968; Wetmur, 1991)

$$k_2 = \frac{k'_N\sqrt{L}}{N} \tag{8-28}$$

Here k'_N is the nucleation rate constant, L is the length of the shortest strand participating in duplex formation, and N is the complexity of the sequence. Complexity has been defined as "the total number of base pairs present in nonrepeating sequences" (Wetmur, 1991) or "the length of DNA needed to contain one copy of the entire sequence" (Cantor

and Schimmel, 1980). The rate constant k'_N is a complicated function of temperature and other conditions, approaching zero at the T_m of a large DNA. It is maximal roughly 25°C below the T_m, where a typical value is $3.5 \times 10^5\ M^{-1}\ s^{-1}$ in 1 M NaCl (Wetmur, 1991). The kinetics of dissociation near the T_m for large DNAs occurs in the 100–1000 s time range and cannot be fit with a simple functional form. This observation is thought to be due to large frictional effects in the unwinding process and is avoided *in vivo* by the actions of nicking-closing enzymes (topoisomerases) and gyrases that break phosphodiester bonds to allow easier rotation.

4.3 Nucleic Acid Hybridization

Nucleic acid hybridization and methods that rely on it are the methods of choice for detecting specific sequences in complex mixtures. For example, the polymerase chain reaction (PCR) and various blotting techniques rely on specific binding of short or long nucleic acid primers or probes to target sequences (Wetmur, 1991). Oligonucleotide arrays on solid support can detect thousands of sequences simultaneously (Schera et al., 1995; Chee et al., 1996). The best conditions for specificity are dependent on many factors, including the T_m and kinetics for the binding and for the disruption of structure in individual sequences. The principles described above can often be used to predict optimum conditions for experiments requiring hybridization (Wetmur, 1991; Steger, 1994). The effects of solid supports, however, have not been studied in detail (Wetmur, 1991).

In applying the principles described above, it is important to recognize the parameter that is important for a given experiment. For example, PCR relies on dissociation of helixes, so the system should be brought to a temperature above T_m to insure complete dissociation. For the annealing step, specificity is important, so the temperature should be a little, typically 10°C, below T_m. For probing blots, the kinetics of dissociation is the important parameter, since unhybridized probe is removed by washing for a given time period. For this application, Wetmur (Wetmur, 1991) suggests defining a dissociation temperature T_d as the temperature at which one-half of the correctly matched hybrid is released in a time t_{wash}. The parameter T_d can be predicted from the expected activation energy for probe dissociation, $E_{A,off}$, the melting temperature T_m (in K) for the hybrid duplex at the probe concentration, and the half-life for probe dissociation at the T_m, $t_{1/2}$:

$$\ln(t_{wash}/t_{1/2}) = (E_{A,off}/R)(T_d^{-1} - T_m^{-1}) \qquad (8\text{-}29)$$

4.4 Tertiary Interactions

The term "tertiary interaction" is used in several different ways. Sometimes it refers to noncanonical interactions, for example a hydrogen bond to a phosphate group as sometimes seen in hairpin and other loops (Varani et al., 1991; Heus and Pardi, 1991; SantaLucia et al., 1992; Pley et al., 1994). Here we use a more restricted definition (Chastain and Tinoco, 1991). If two nucleotides i and j are paired, then a tertiary interaction is any interaction, direct or indirect, between a nucleotide k between i and j in the sequence, and another nucleotide ℓ which is not between i and j. Thus a pseudoknot (Fig. 8-11) is one form of tertiary interaction. Another is metal ion mediated interactions between nucleotides k and ℓ (Lu and Draper, 1994).

One approach to understanding tertiary interactions is to study binding of oligonucleotides to natural RNAs.The methods employed include kinetics (Sugimoto et al., 1988, 1989b; Herschlag and Cech, 1990; Bevilacqua et al., 1992), fluorescence (Bevilacqua et al., 1992; Sugimoto et al., 1989a), equilibrium dialysis (Bevilacqua and Turner, 1991), and gel retardation (Pyle et al., 1990). The last method is the fastest, but is not completely understood. When using gel retardation, it is important to allow sufficient incubation time before loading samples on the gel (Bevilacqua and Turner, 1991; Pyle et al., 1994). Another method for studying tertiary interactions is to monitor the effects of site directed mutation on RNA folding and function (Murphy and Cech, 1994; Costa and Michel, 1995).

Little is known about the thermodynamics of tertiary interactions. One interaction that has been identified involves hydrogen bonding to 2'-OH groups for recognition of an oligomer substrate by a group I ribozyme (Sugimoto et al., 1989b; Pyle and Cech, 1991; Bevilaqua and Turner, 1991; Strobel and Cech, 1993). Each such interaction can provide a free energy increment of about 1 kcal mol^{-1} (Bevilacqua and Turner, 1991). In one case, the interaction involves a hydrogen bond from the H of the 2'-OH to the N1 of an A (Pyle et al., 1992).

Unfavorable tertiary interactions are also possible. For example, UCdGU binds to a group I ribozyme at least 30-fold more weakly than UCdG (Moran et al., 1993). A model of the binding site (Michel and Westhof, 1990) suggests that it is designed to put strain on the phosphodiester bond between G and U. This bond is the site of cleavage in the second step of group I splicing, and such strain may provide a catalytic advantage (Moran et al., 1993). Such substrate destabilization has also been reported in a ribozyme model system for the first step of splicing (Bevilacqua et al., 1994; Narlikar et al., 1995).

Favorable and unfavorable tertiary interactions are presumably sensitive to the overall folding of an RNA. Thus they can give rise to cooperativity and anticooperativity in binding of substrates. Such effects have also been observed with a group I ribozyme (Bevilacqua et al., 1993; McConnell et al., 1993).

5. ENVIRONMENTAL EFFECTS ON HELIX STABILITY

5.1 Salt Concentration

A theoretical treatment of salt effects on helix stability for polyelectrolytes is given in Chapter 11. Here we focus on experimental results. In solutions containing only Na^+ or similar ions, increasing salt concentration up to about 1 M continuously increases helix stability. Up to about 0.2 M Na^+, the increase in T_m is linear with log [Na^+] (see Eq. 11-20). The rate of increase depends on base composition. For example, from fitting data obtained on different, natural DNAs (Owen et al., 1969), Frank-Kamenetskii (1971) found $dT_m/d\log[Na^+] = 18.30 - 7.04F_{GC}$ (F_{GC} is the fractional GC content). From studies of subtransitions in the melting of lambda DNA, Blake and Haydock (1979)

proposed a similar equation: $dT_m/d\log[Na^+] = 19.96 - 6.65F_{GC}$. Thus, the T_m of a deoxypolynucleotide with only AT base pairs increases almost 20°C for every factor of 10 increase in $[Na^+]$.

At salt concentrations above 1 *M*, addition of salt lowers the T_m of DNA. The lowering is relatively independent of cation, but is strongly dependent on anion with $CCl_3COO^- > SCN^- > ClO_4^- > CH_3COO^- > Br^-, Cl^-$ (Hamaguchi and Geiduschek, 1962). This order correlates with the effect of these ions on the solubility of the bases (Robinson and Grant, 1966). The better denaturants are most effective in enhancing base solubility.

There are few studies of the effect of Mg^{2+} concentration on polynucleotide melting (Riesner and Römer, 1973; Blagoi et al., 1978; Krakauer, 1974). As with Na^+, addition of Mg^{2+} initially increases T_m. The effect saturates at about $10^{-3} - 10^{-2}M$, however. Thereafter, addition of Mg^{2+} lowers T_m (Blagoi et al., 1978). Somewhat counterintuitively, addition of Na^+ in the presence of enough Mg^{2+} to neutralize the backbone charges results in a lowering of T_m. This lowering of T_m is expected from theoretical considerations (Record, 1975; Manning, 1978). In particular, at saturating concentrations of Mg^{2+}, the effective charge on single strands is larger than on double strands due to counterion condensation. Thus, single strands are stabilized more by increases in ionic strength because of Debye–Hückel screening effects.

Relatively little is known about salt effects on stabilities of oligonucleotide helixes. Qualitatively, the effects are similar to those observed for polynucleotides. For example, $dT_m/d\log[Na^+]$ for dGCATGC and dGGAATTCC are 11 and 12°C, respectively (Williams et al., 1989; Erie et al., 1987), only a little less than predicted for polymers with the same GC content. This result is somewhat surprising since theoretical considerations indicate less charge will be neutralized in oligonucleotides than in polynucleotides because of the reduced charge density at the ends (Record and Lohman, 1978; Olmsted et al., 1989). Relatively little experimental data is available, however. One interesting question is: At what length does an oligomer behave much like a polymer? This question has not been investigated in detail. The values of $dT_m/d\log[Na^+]$ for A_7U_7 and poly (A) poly (U) are similar, however: 17.4 and 19.6°C, respectively (Hickey and Turner, 1985b).

5.2 Solvent Effects

Addition of cosolvents to aqueous solutions of nucleic acids typically destabilizes the ordered form, and this is often used as a substitute for denaturation by temperature (Lerman et al., 1984). Typically, the T_m of a double helix will be a linear function of cosolvent concentration (Klump and Burkart, 1977; Hickey and Turner, 1985b). For example, the T_m values of calf thymus and salmon sperm DNA decrease 2.5°C M^{-1} urea (Klump and Burkart, 1977). The effects of cosolvents have been studied by measuring the concentration of cosolvent required to give 50% denaturation of T4 DNA at 73°C (Levine et al., 1963) and by measuring the effect of cosolvents on thermal denaturation curves of various oligonucleotides (Hickey and Turner, 1985b; Albergo and Turner, 1981). Some of these results are listed in Table 8-15. Somewhat surprisingly, there is no simple correlation between the concentration required for 50% denaturation at 73°C and the effect on T_m. This finding may reflect the very different conditions and assays

Table 8.15
Effects of Denaturants on Duplex Stability

Cosolvent	Molarity for 50% denaturation of T4 DNA at 73°C, 43 mM Ionic Strength		Δt_m(°C) at 10 mol % for 18.7μM Oligo in 1 *M* NaCl	
	Predicted[a]	Observed[b]	A_7U_7[c]	$(dGC)_3$[d]
Alcohols				
Methanol	3.9	3.5		6.6
Ethanol	1.3	1.2	8.3	11.1
1-Propanol	0.47	0.54	9.1	8.4
2-Propanol	0.64	0.90	8.4	
Ethyleneglycol	1.7	2.2	7.7	
Glycerol		1.8	7.9	
Cyclohexyl alcohol		0.22		
Phenol		0.08		
Other Compounds				
Pyridine		0.09		
1,4 Dioxane		0.64	18.7	
Formamide	1.5	1.9	14.8	12.0
N, *N* dimethylformamide	0.54	0.60	16.9	~22
Urea	1.1	1.0	17.8	13
Acetonitrile		1.2		
TritonX-100		> 10%		

[a]Herskovits and Harrington, (1972); Herskovits and Bowen, (1974).

[b]Levine et al., (1963).

[c]Hickey and Turner, (1985).

[d]Albergo and Turner, (1981).

used in the experiments. For example, the cosolvent required for 50% denaturation was measured for an ionic strength of 0.04 with an antibody assay, whereas the effect on T_m was measured at 1 *M* Na^+ with optical melting curves. The cosolvent concentrations required for 50% denaturation correlate well with enhancement of base solubility (Levine et al., 1963; Herskovits and Harrington, 1972; Herskovits and Bowen, 1974), and this has predictive value (see Table 8-15). This correlation seems reasonable since the bases are more exposed to solvent in the denatured form, and favorable interactions between bases and cosolvent favor denaturation.

5.3 pH

The stability of most double helixes is relatively insensitive to pH between 5 and 9. At lower and higher pH values, stability is decreased (Bloomfield et al., 1974; Blake, 1996). At low pH, bases in the single strand bind more protons than in the duplex, thus favoring single strand. At high pH, guanine, thymine, and uracil are deprotonated, thus precluding normal hydrogen bonding and increasing charge repulsion. At both

extremes, the charge on the single strand is different from that around pH 7. This affects the salt dependence of the T_m (Manning, 1978; Record et al., 1976). In particular, $dT_m/d\log[Na^+]$ will be larger at low pH and smaller at high pH. For example, $dT_m/d\log[Na^+] \approx 0$ near pH 11 (Record, 1967).

Some sequences form double helixes only at low pH. For example, poly(A), poly(dA), poly(C), and poly(dC) form duplexes with transition midpoints between pH 4 and 8 depending on conditions of salt and temperature (Inman, 1964; Hartman and Rich, 1965; Holcomb and Timascheff, 1968; Adler et al., 1969; Guschlbauer, 1975). This results from hydrogen bonding schemes in which a proton is shared between two bases, for example $A \cdot A^+$. The protonated pairs do not have to be adjacent. Poly(dCT) forms a duplex with $C \cdot C^+$ pairs alternating with unpaired thymines (Gray et al., 1980). Duplex formation by $C \cdot C^+$ pairing has also been observed at the ends, but not the middle, of mixed deoxy sequences. For example, $d(C_4A_4T_4C_4)$ forms such duplexes (Gray et al., 1984), but $d(A_{10}C_4T_{10})$, $d(AACC)_5 + d(CCTT)_5$, and $d(A_6C_6A_6) + d(T_6C_6T_6)$ do not (Edwards et al., 1988). This context dependence indicates $C \cdot C^+$ pairs are hard to form between stretches of AT pairs. In RNA, however, $C \cdot C^+$ pairs have been observed in the middle of the duplex $(CGCCCGCG)_2$ (SantaLucia et al., 1991). Triple-strand helixes involving $C \cdot G \cdot C^+$ pairing have also been observed to form upon protonation of C (Lee et al., 1979). Little is known about the pH dependence of structure for other sequences, but there are suggestions that interesting effects will be discovered (Topping et al., 1988; Kao and Crothers, 1980; Legault and Pardi, 1994). For example, it has been shown that 5S rRNA undergoes a tertiary structural switch between pH 7 and 8 (Kao and Crothers, 1980).

6. TRIPLE HELIXES

Hydrogen bonded base triples are found in tRNA (see Fig. 4.13). A triple-stranded structure consisting of $T \cdot AT$ and $C^+ \cdot GC$ base triples has been proposed for regions of DNA containing $(dT\text{-}dC)_n \cdot (dA\text{-}dG)_n$ repeats (Christophe et al., 1985; Mirkin et al., 1987; Voloshin et al., 1988; Htun and Dahlberg, 1988; Johnston, 1988; Wells et al., 1988; Mirkin and Frank-Kamenetskii, 1994). When this triple strand forms, the remaining $(dA\text{-}dG)_n$ strand is released from pairing as shown in Figure 8-14, and the overall structure is referred to as H-DNA. It has also been proposed that pairing to give triple strands could regulate transcription (Cooney et al., 1988; Miller and Sobell, 1966) and be used for site specific modification or cleavage of nucleic acid (Strobel et al., 1988; Praseuth et al., 1988). These ideas may provide the basis for therapeutics that bind as a third strand.

Only sequences that are largely polypurine and polypyrimidine have been observed to form triple strands (Lee et al., 1979; Wells et al., 1988; Felsenfeld and Miles, 1967; Michelson et al., 1967). Known triple-strand pairings include: $U \cdot AU$ (Felsenfeld et al., 1957; Stevens and Felsenfeld, 1964), $A \cdot UA$ (Broitman et al., 1987), $T \cdot AT$ (Riley et al., 1966), $C^+ \cdot GC$ (Lee et al., 1979; Lipsett, 1964), $T \cdot CG$ (Yoon et al., 1992), $U \cdot CG$ (Michel et al., 1990; Michel and Westhof, 1990), $G \cdot TA$ (Yoon et al., 1992; Griffin and Dervan, 1989), $G \cdot CG$ (Lipsett, 1964; Lipsett, 1963), and $A \cdot GC$ (Chastain and Tinoco, 1992). Triple-strand helixes have also been observed with 2′-5′ linked oligonucleotides (Jin et al., 1993). The most thoroughly studied triplex is $poly(A) \cdot 2$

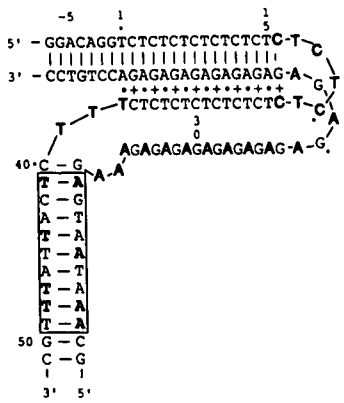

Figure 8-13
Schematic of structure for H-DNA [Reprinted with permission from Htun, H. and Dahlberg, J. E. (1988). *Science* **241**, 1791–1796. Copyright ©American Association for the Advancement of Science.] In this structure, Watson–Crick base pairs break to allow formation of a triple helix.

poly(U) (Felsenfeld et al., 1957; Stevens and Felsenfeld, 1964; Ross and Scruggs, 1965; Krakauer and Sturtevant, 1968). The stoichiometry of this complex was obtained from optical mixing curves (Felsenfeld et al., 1957; Stevens and Felsenfeld, 1964). The stability of the triplex relative to duplex and single strands depends on temperature and salt concentration as shown in the phase diagram in Figure 8-15 (Stevens and Felsenfeld, 1964; Krakauer and Sturtevant, 1968). As expected, high-salt favors triplex which is true for Mg^{2+} as well as for Na^+ (Felsenfeld et al., 1957). For example, in region 2 in Figure 8-15 the combination of $poly(A) \cdot 2\ poly(U)$ + poly(A) is more stable than 2 [poly(A)·poly(U)]. Thus any duplexes present will disproportionate to form triplex and single-strand poly(A):

$$2[\text{poly(A)poly(U)}] \rightleftharpoons [\text{poly(A)2 poly(U)}] + \text{poly(A)}$$

The $\Delta H°$ for this reaction is about 4 kcal mol^{-1} (Krakauer and Sturtevant, 1968). The $\Delta H°$ for formation of triplex from duplex and poly(U) is about −4 kcal mol^{-1} (Ross and Scruggs, 1965; Krakauer and Sturtevant, 1968). The time required for triple helix formation ranges from less than 1 to many minutes depending on length and conditions (Felsenfeld et al., 1957; Pörschke and Eigen, 1971).

7. G-QUARTETS

Hydrogen-bonded *G* quartets are known to occur in the telomeric regions at the ends of chromosomes (Williamson et al., 1989), and G tetraplexes can form from G monomers (Gellert et al., 1962; Pinnavaia et al., 1978), G rich DNA and RNA oligomers (Sen and Gilbert, 1988; Sundquist and Klug, 1989; Sen and Gilbert, 1990; Kim et al., 1991;

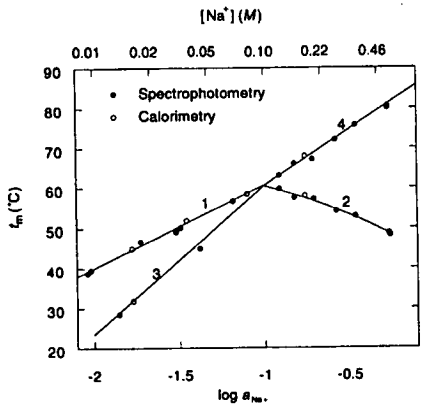

Figure 8-14
Dependence of t_m on the concentration of Na^+ at neutral pH in the absence of divalent ions for various reactions of polyA and polyU [Heats of Helix–Coil Transactions of PolyA- PolyU Complexes, Krakauer, H. and Sturtevant, J. M. *Biopolymers*, 6, 491–512. Copyright ©1968. Reprinted by permission of John Wiley & Sons, Inc.] The lines correspond to the following reactions: (1) poly(A·U) → poly(A) + poly(U), (2) poly(A·U) → 1/2 poly(U·A·U) + 1/2 poly(A), (3) poly(U·A·U) → poly (A·U) + poly(U), (4) poly(U·A·U) → poly(A) + 2 poly(U). Note that the topmost region corresponds to single stranded poly(A) + poly(U).

Hardin et al., 1991; Jin et al., 1992; Lu et al., 1993), and polymers (Zimmerman et al., 1975). Formation of the G tetrad is favored by K^+ relative to Na^+ (Gellert et al., 1962; Pinnavaia et al., 1978; Sen and Gilbert, 1990; Hardin et al., 1991; Hud et al., 1996). It has been suggested that formation of the G tetrad may be important for several biological processes (Williamson et al., 1989; Sen and Gilbert, 1988; Sundquist and Klug, 1989; Sen and Gilbert, 1990; Williamson, 1994). The thermodynamics of tetraplex formation has been studied for several oligomer sequences (Jin et al., 1992; Lu et al., 1993). In the presence of K^+, ΔG°_{25} and ΔH° per tetrad appear to be -2 to -3 kcal mol^{-1} and -20 to -30 kcal mol^{-1}, respectively (Jin et al., 1992).

8. PREDICTING SECONDARY STRUCTURE

A major application of our knowledge of conformational changes is the prediction of nucleic acid structure from sequence. Most applications have involved RNA, but the methods are also applicable to DNA. One goal of these methods is to predict the base pairing, or secondary structure, for single-stranded chains. Experiments indicate that tertiary structure is less stable than most secondary structure (Crothers et al., 1974; Banerjee et al., 1993; Jaeger et al., 1993; Laing and Draper, 1994). Thus as a first approximation, tertiary interactions can be neglected when predicting secondary structure. The conclusion that stabilities of small nucleic acid structures are largely determined by strong, local interactions such as stacking and hydrogen bonding suggests that summing the free energy changes for such interactions will provide an approximation for the stability of a given structure (Turner et al., 1988; Papanicolaou et al., 1984; Tinoco et al., 1971). The structure predicted to be most prevalent at equilibrium is then the one with the lowest free energy (see Eqs. 8-3 and 8-7a). It should be realized, however, that the relative concentrations predicted for various species is quite sensitive to the ΔG° values in the calculation. The predicted equilibrium constant between two species at 37°C changes by a factor of 10 for every 1.4 kcal mol^{-1} in ΔG°. Thus current predictions should be considered rough approximations.

As discussed in Section 2, stabilities of oligonucleotides without loops can be predicted well with a nearest neighbor model. Thus, this approximation provides a reasonable treatment for helical regions. Much less is known about the sequence dependence of stability for structures with loops, and current methods largely neglect this sequence dependence. This restriction can be eliminated when more experimental data become available. Even with these limitations, the loop parameters listed in Table 8-14 are useful for predicting RNA secondary structure.

If sequence dependent free energy parameters were available for every possible loop, it would still be difficult to completely predict the most stable conformation from sequence. The reason is that the time required to try every possibility is usually enormous. For example, for a sequence of N nucleotides with A, C, G, and U occurring randomly, the number of valid structures goes approximately as 1.8^N (Zuker and Sankoff, 1984). If ΔG° could be calculated for 1000 structures every second, it would take about 10^{10} years, roughly the age of the universe to try all valid possibilities for an 80 nucleotide sequence. To circumvent this problem, clever computer algorithms have been written that avoid trying every possibility (Papanicolaou et al., 1984; Zuker and Stiegler, 1981; Williams and Tinoco, 1986; Nussinov et al., 1982; Zuker, 1989; Mathews et al., 1999; Rivas and Eddy, 1999). These algorithms, however, require additional approximations. Two of the approaches are discussed below.

8.1 Combinatorial Algorithms

Combinatorial algorithms develop a list of all helixes that can be formed from a sequence (Papanicolaou et al., 1984; Gouy, 1986). These algorithms can include knotted structures. The algorithms then try combinations of these helixes in search of the lowest free energy. Various tricks are used to avoid computing combinations that are not likely to be low in free energy. Nevertheless, the time required is large because the number of helixes, L, and the number of combinations grow approximately as N^2 and 2^L, respectively.

8.2 Recursive Algorithms

Recursive (or dynamic) algorithms usually make the approximation that if two nucleotides pair, then any nucleotide between them in the primary sequence can only pair

with other nucleotides between the originally paired nucleotides (Zuker and Sankoff, 1984; Williams and Tinoco, 1986; Nussinov et al., 1982). Thus knots are usually not allowed, but the other loops shown in Figure 8-11 are allowed. This approximation permits the lowest free energy structure to be found from consideration of the lowest free energy structure for each possible subfragment of the sequence. The computation for each new subfragment makes use of the computations for each smaller subfragment, which makes the algorithms recursive, and therefore fast. The time required typically grows as N^3 or N^4, depending on the generality of the free energy parameters. When knots are included in a recursive algorithm, the time grows as N^6 (Rivas and Eddy, 1999).

The most popular recursive algorithm can use more than thermodynamic parameters for deducing structure (Zuker and Stiegler, 1981; Zuker, 1989; Mathews et al., 1999). For example, if it is known that a given nucleotide is susceptible to nuclease cleavage, the predicted structure can be forced to include this. Thus experimental data can play a role in the process.

8.3 Suboptimal Structures

Except for short sequences, free energy minimization may never lead reliably to the exact secondary structure, because many approximations are mandated by experimental and computational considerations. Even if it were possible to predict the most stable structure, it is likely that other structures will be important for understanding function. For these reasons, both combinatorial and recursive algorithms have been designed to permit prediction of suboptimal structures (Steger et al., 1984; Williams and Tinoco, 1986; Zuker, 1989; Gouy, 1986). The output from these programs can then be tested against experimental data (Mathews et al., 1997) or used to design experiments aimed at testing various possibilities. If more than one sequence is known for molecules with similar functions, then comparison of suboptimal structures may help identify the true structure and features required for function (Lück et al., 1996; Mathews et al., 1997). Figure 8-16 provides an example of a sequence with two structures that are similar in free energy. Programs for folding RNA are now available on the WEB (e.g., http://rna.chem.rochester.edu, and http://www.ibc.wustl.edu/~zuker).

8.4 Evaluating Predictions

The performance of free energy minimization procedures can be tested by comparing predicted structures for various RNA sequences with those deduced from sequence comparisons. In tests of a recursive algorithm on various RNA sequences, roughly 70% of the known base pairs are present in the free energy minimized structures (Mathews et al., 1999). The best computer-generated suboptimal structure has roughly 85% of the known base pairs (Mathews et al., 1999). Considering the lack of experimental data for many free energy parameters, the required approximations, and the difficulty of the problem, the agreement between known and energy minimized structures is surprisingly good. This finding supports the assumption that secondary structure is largely determined by local interactions and is not very dependent on tertiary structure.

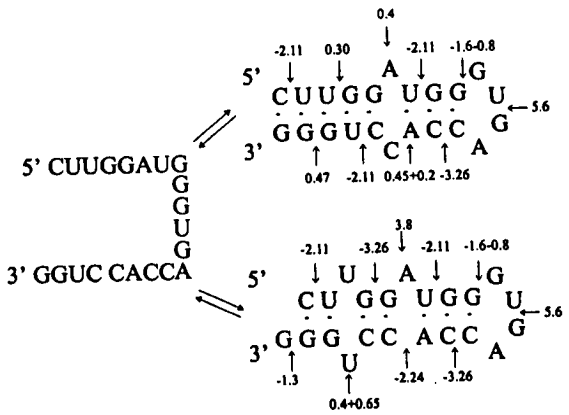

Figure 8-15
Example of an RNA sequence with two secondary structures of similar free energy. The sequence is a subfragment of 5S rRNA from *Philosamia cynthia ricini*. A program that gives suboptimal structures will predict both structures.

Presumably, performance will continue to improve as more parameters are measured experimentally.

It may also be possible to improve predictions of structure from sequence by adding considerations from the kinetics of folding. From the forward and reverse rate constants of hairpin formation, it is clear that a nucleic acid is similar to a computer algorithm in that it does not have time to try all possible pairings. Thus the pathway of folding must also be coded into the sequence. So far, only the folding of tRNA has been studied in detail (Riesner and Römer, 1973; Crothers and Cole, 1978; Crothers et al., 1974). Therefore general rules for folding are not clear yet. Preliminary attempts have been made, however, to introduce kinetics into algorithms for predicting secondary structure (Gultyaev et al., 1995; Schmitz and Steger, 1996).

9. PREDICTING THREE DIMENSIONAL STRUCTURE

Once the secondary structure of a nucleic acid is known, the next challenge is to predict the 3D structure. Good models are available for the 3D structures of double helixes (see Chapter 4), and determination of all the double helical regions greatly constrains possible 3D foldings. It is still necessary, however, to determine folding between helical regions. Presumably, this will be determined by the same factors that are important for

secondary structure: conformational entropy, stacking, hydrogen bonding, counterion binding, and perhaps specific solvation. Contributions from all these factors can be seen in the crystal structures of tRNA and part of a group I intron (see Figs. 4.11–4.14 and color plate for 8-17) (Kim et al., 1974; Robertus et al., 1974; Cate et al., 1996). The limited data base available, however, makes it difficult to deduce general principles.

Since the factors governing 3D and secondary structure seem similar, it should be possible to use information from studies of conformational changes in oligonucleotides and polynucleotides to help predict 3D structure. The structure of yeast phenylalanine tRNA shown in Figure 8-17 illustrates three correlations of 3D structure with free energy changes measured for stacking in oligonucleotides. The first correlation involves single-strand stacking. Table 8-1 indicates that single-strand AA and CC sequences have considerable stacking. In the tRNA crystal structure, the sequences 35AA36 and 74CC75 are both stacked. Conversely, Table 8-1 indicates single-strand UU sequences do not stack. Dihydrouracil, D, is similar to uracil, except the pyrimidine ring is not planar. This lack of planarity should lead to even less tendency to stack. Inspection of Figure 8-17 shows that both dihydrouracils in the sequence 16DD17 are unstacked.

The second correlation involves coaxial stacking of adjacent helixes. In the tRNA crystal structure, the base pairs 7UA66 and 49CG65 are stacked on each other. This type of interaction is also associated with a large favorable free energy change in an oligonucleotide model system (Walter et al., 1994; Walter and Turner, 1994).

The third correlation is observed for stacking of unpaired nucleotides (dangling ends) adjacent to base pairs (Sugimoto et al., 1987a; Turner et al., 1988; Burkard et al., 1999). Figure 8-17 (*a*) illustrates dangling end sequences in yeast phenylalanine tRNA that have ΔG° values more favorable than -1 kcal mol in Table 8-5. These are strongly stacking sequences in oligonucleotides, and in the tRNA structure each such dangling end is stacked on its adjacent base pair. Figure 8-17 (*b*) illustrates dangling end sequences in tRNA that have ΔG° values less favorable than -0.4 kcal mol^{-1}. These weakly stacking sequences are not stacked on the adjacent base pairs in tRNA. In four out of five such cases, these weakly stacking sequences occur at places where there is a turn in the sugar-phosphate backbone. These are the only places in the structure where sharp turns occur at the ends of helical regions, which suggests weakly stacking sequences are favored at positions where the backbone turns. While limited, the results suggest free energies measured in oligonucleotides reflect the strengths of fundamental interactions that will help determine 3D as well as secondary structure.

Studies of oligonucleotides binding to natural RNAs are suggesting additional interactions that will be important for determining 3D structure. For example, interactions with 2′-OH groups, can be used for aligning helixes (Sugimoto et al., 1989; Pyle and Cech, 1991; Bevilacqua and Turner, 1991; Pyle et al., 1992). This type of helix alignment was suggested by a phylogenetic analysis of about 80 sequences (Michel and Westhof, 1990). This analysis also suggested several other tertiary interactions such as triple helix formation and interactions with tetraloops. Some have been confirmed by site directed mutagenesis (Michel et al., 1990; Costa and Michel, 1995), and by X-ray diffraction (Pley et al., 1994; Cate et al., 1996). Presumably, as more structures are determined, more patterns will be recognized. For example, pairing between complementary loops could also be important.

10. HELIX-HELIX TRANSITIONS

This chapter has focused on helix–coil transitions. Transitions are also known in which a helix changes conformation. Transitions between A, B, and Z forms have been studied. The Z form is most stable for alternating CG sequences, but variations are possible (Jovin and Soumpasis, 1987). The A and B forms can be stabilized for random sequence deoxy polymers, but the B form has not been observed for polyribonucleotides. Table 8-16 lists conditions under which various conformations are stable. Often unusual conformations are induced by conditions that affect the effective salt concentration and water activity. For example, A → Z, B → Z, and B → A transitions are all promoted by ethanol. High salt promotes A and B → Z transitions. Another similarity is that these transitions typically have small ΔH° values that are often positive (Klump and Jovin, 1987; Chaires and Sturtevant, 1986). Thus, for example, high temperature tends to promote the Z conformation. Despite extensive characterization of these transitions, there is no consensus on the quantitative contributions of various fundamental interactions. It is likely, however, that charge, solvation, and steric effects are important. There is also no clearly documented case in which these types of transitions serve a known physiological function. Discovery of functional significance is probably only a matter of time.

APPENDIX

A.1 Statistical thermodynamics of transitions

The two-state model for deriving thermodynamic parameters from spectroscopic data is not general. While short oligomers often exhibit two-state transitions, long oligomers

Table 8.16
Conditions Favoring Unusual Conformations

Transition	Reference	Cation or Salt	Solvent	T(°C)
A → Z	Cruz et al., (1986)	4.8–6 *M* $NaClO_4$	H_2O	35–45
[Poly(rCrG)]	Cruz et al., (1986)	4.8 *M* $NaClO_4$	20% EtOH	25
	Cruz et al., (1986)	5–7 *M* NaBr	H_2O	~20–65
	Trulson et al., (1987)	3.8 *M* $MgCl_2$	H_2O	~20
B → Z	Pohl and Jovin (1972)	4 *M* NaCl	H_2O	25
[Poly(dCdG)]	Pohl and Jovin (1972)	2.5 *M* $NaClO_4$	H_2O	25
	Pohl (1976); Hall and Maestre (1984)	~ 5 × 10^{-4} *M* Na^+	10-50% EtOH	20-50
	Pohl and Jovin (1972)	1 *M* $MgCl_2$	H_2O	25
	Behe and Felsenfeld (1981)	2 × 10^{-5} *M* $Co(NH_3)_6^{3+}$	H_2O	20
B → Z	Behe and Felsenfeld (1981)	≥ 0.7 *M* NaCl	H_2O	~20
[Poly(dm5CdG)]	Behe and Felsenfeld (1981)	≥ 6 × 10^{-4} *M* $MgCl_2$	H_2O	~20
B → A, [DNA]	Ivanov et al. (1974)	~ 5 × 10^{-4} *M* NaCl	80% EtOH	5–30

do not. Often terminal base pairs, base pairs adjacent to loops, or regions of weak base pairs melt first. Analysis of these non-two-state transitions requires a statistical thermodynamic model (Wada et al., 1980; Gotoh, 1983; Wartell and Benight, 1985; Bloomfield et al., 1974; Cantor and Schimmel, 1980; Poland, 1978).

In any statistical thermodynamic model, all the thermodynamic information is contained in the molecular partition function, q:

$$q = \sum_{j=0}^{n} \exp[-G_j/RT] \tag{8-A.1}$$

where G_j is the free energy of the jth configuration, and the summation is over all possible configurations. When only duplex formation is being considered, G_j for the completely single stranded state is set at 0, and the remaining G_j values are replaced by ΔG_i values, the difference in free energy between a given duplex configuration and the single-stranded state, to give

$$q = 1 + \sum_{i=1}^{n} \exp[-\Delta G_i/RT] \tag{8-A.2}$$

It is more convenient, however, to work with the conformational partition function, q_c:

$$q_c = q - 1 = \sum_{i=1}^{n} \exp[-\Delta G_i/RT] = \sum_{i=1}^{n} K_i \tag{8-A.3}$$

where K_i is the equilibrium constant for forming duplex configuration i from single strands. Any model can now be used to derive an expression for q_c. The trick is to use a model complicated enough to describe the experimental results, but simple enough to permit unambiguous extraction of the necessary parameters from data. This approach will be illustrated with the simplest model, the zipper model (Bloomfield et al., 1974; Cantor and Schimmel, 1980; Poland, 1978).

In the zipper model, only one helical region is allowed per duplex. To illustrate the process in its simplest form, we assume that the equilibrium constant for adding one base pair to an existing helix is always the same, s; that the equilibrium constant for initiating the helix is κ (often labeled σs in the literature); and that only perfectly aligned helixes are stable enough to contribute to q_c. These approximations are adequate for a sequence such as $G_3C + GC_3$ (see Figure 8-18). With these assumptions, the equilibrium constant for forming each duplex configuration with j base pairs from the single strands is κs^{j-1}. If there are $g_j(N)$ distinguishable duplexes with j base pairs that can be formed by a single strand with N bases, then Eq. 8-A.3 can be written

$$q_c = \kappa \sum_{j=1}^{N} g_j(N) s^{j-1} \tag{8-A.4}$$

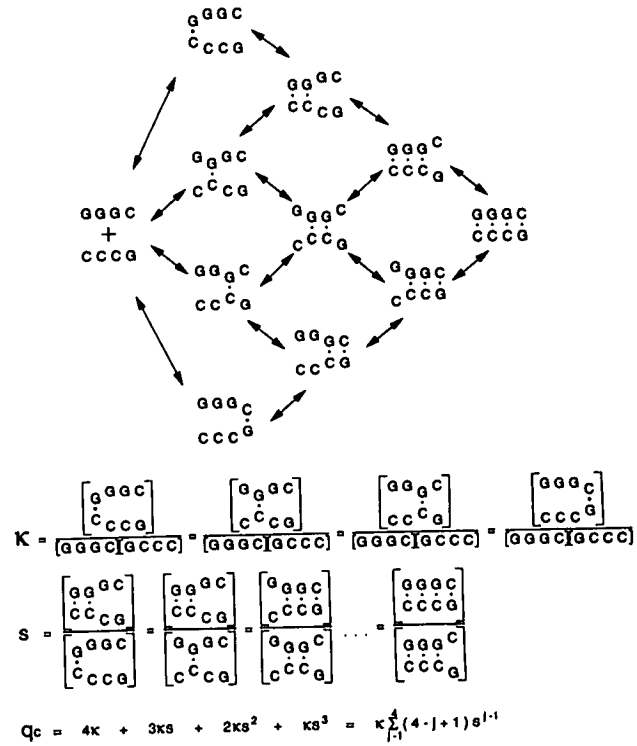

Figure 8-18
Application of the zipper model to calculate the conformational partition function, q_c, for duplex formation by G_3C+GC_3. The zipper model allows only one helical region per duplex.

When the two strands in the duplex have different sequence (i.e., are nonself-complementary), and only completely aligned duplexes are allowed (see Fig. 8-18), then $g_j(N) = N - j + 1$, giving

$$q_c = \kappa \sum_{j=1}^{N} (N - j + 1) s^{j-1} = \kappa[(N+1) \sum_{j=1}^{N} s^{j-1} - \sum_{j=1}^{N} j s^{j-1}] \tag{8-A.5}$$

These are related to the finite geometric series, for which it can be shown that

$$\sum_{j=1}^{N} s^{j-1} = \frac{s^N - 1}{s - 1} \tag{8-A.6}$$

$$\sum_{j=1}^{N} j s^{j-1} = \frac{N s^{N+1} - (N+1)s^N + 1}{(s-1)^2} \tag{8-A.7}$$

so

$$q_c = \kappa \left(\frac{(N+1)(s^N - 1)}{s - 1} - \frac{N s^{N+1} - (N+1)s^N + 1}{(s-1)^2} \right) = \kappa \frac{s^{N+1} - (N+1)s + N}{(s-1)^2} \tag{8-A.8}$$

Note that if $s > 1$ and N is large, $q_c \approx \kappa s^{N-1}$. Thus, in this limit q_c is simply the equilibrium constant for the two-state model.

To determine q_c it is necessary to measure κ and s. One way to do this is by analyzing optical melting curves. In the simplest case, the concentrations of the two complementary strands are equal, and the absorbance A of a sample depends only on the number of base pairs in the sample. For this case, the fraction X_b of bases paired is given by

$$X_b = \frac{A - A_s}{A_d - A_s} \tag{8-A.9}$$

Here A_s and A_d are absorbances when the sample is completely single strands or fully paired duplexes, respectively. (Note the similarity to Eq. 8-14, with α corresponding to X_b.) We need to express X_b in terms of q_c. For nonself-complementary strands, A and B, the equilibrium can be written

$$A + B \rightleftharpoons C_1 + C_2 + \cdots + C_n \tag{8-A.10}$$

where $C_1, C_2, \ldots, C_n$ are the possible configurations of duplex (see Fig. 8-18). Assuming the total concentrations of strands A and B are equal.,

$$q_c = \frac{[C_1] + [C_2] + \ldots + [C_n]}{[A][B]} = \frac{0.5XC_T}{[0.5(1 - X)C_T]^2} = \frac{2X}{(1 - X)^2 C_T} \tag{8-A.11}$$

where C_T is the total concentration of strands,

$$C_T = [A] + [B] + 2\sum_{i=1}^{n} [C_i] \tag{8-A.12}$$

and X is the fraction of strands in duplex,

$$X = 2([C_1] + [C_2] + \cdots + [C_n])/C_T \tag{8-A.13}$$

The average number of base pairs per duplex, $< n >$, can be calculated by realizing that the fraction of duplexes with j base pairs, f_j, is equal to $\kappa g_j(N) s^{j-1}$, normalized by the sum over all j, q_c. Thus

$$< n > = \sum_{j=1}^{N} j f_j = \frac{1}{q_c} \sum_{j=1}^{N} j \kappa g_j(N) s^{j-1} = \frac{1}{q_c} \frac{d}{ds} \sum_{j=1}^{N} \kappa g_j(N) s^j = \frac{1}{q_c} \frac{d(sq_c)}{ds} \tag{8-A.14}$$

This equation allows X_b to be expressed in terms of X and q_c:

$$X_b = \frac{X < n >}{N} = \frac{X}{N q_c} \frac{d(sq_c)}{ds} \tag{8-A.15}$$

Solving Eq. 8-A.11 for X in terms of q_c leads to

$$X_b = \frac{1 + q_c C_T - \sqrt{1 + 2q_c C_T}}{N q_c^2 C_T} \frac{d(sq_c)}{ds} \tag{8-A.16}$$

By using Eq. 8-A.9, X_b can be measured as a function of C_T and/or N, and the resulting simultaneous equations solved for κ and s.

While the above example illustrates the main features of the statistical approach, it is not general. For intramolecular helix formation with the same assumptions, the theory is similar (Bloomfield et al., 1974; Cantor and Schimmel, 1980). For most real cases, however, it is more complicated. Most sequences do not have the same s for each base pair (e.g., see Table 8-4). Many duplexes contain loops of various kinds so that more than one helical region is present. In this case, additional helix initiation parameters must be added that are different from κ. In general, the absorbance will not depend only on the number of base pairs (Poland, 1978). Thus X_b in Eq. 8-A.9 must be replaced by a more complicated function. Clearly, the number of parameters increases rapidly for real systems, which explains the relatively rare application of statistical models.

References

Aboul-ela, F., Koh, D., Tinoco, I., Jr., and Martin, F. H. (1985). Base-base Mismatches. Thermodynamics of Double Helix Formation for $dCA_3XA_3G + dCT_3YT_3G$ (X,Y = A,C,G,T). *Nucleic Acids Res.* **13**, 4811–4824.

Adler, A. J., Grossman, L., and Fasman, G. D. (1969). Polyriboadenylic and Polydeoxyriboadenylic Acids. Optical Rotatory Studies of pH-Dependent Conformations and Their Relative Stability. *Biochemistry* **8**, 3846–3859.

Albergo, D. D., Marky, L. A., Breslauer, K. J., and Turner, D. H. (1981). Thermodynamics of $(dG\text{-}dC)_3$ Double-Helix Formation in Water and Deuterium Oxide, *Biochemistry* **20**, 1409–1413.

Albergo, D. D. and Turner, D. H. (1981). Solvent Effects on Thermodynamics of Double-Helix Formation in $(dG\text{-}dC)_3$, *Biochemistry* **20**, 1413–1418.

Allawi, H. T. and SantaLucia, J., Jr. (1997). Thermodynamics and NMR of Internal G·T Mismatches in DNA, *Biochemistry* **36**, 10581–10594.

Allawi, H. T. and SantaLucia, J., Jr. (1998). Nearest Neighbor Thermodynamic Parameters for Internal GA Mismatches, *Biochemistry* **37**, 2170–2179.

Antao, V. P. and Tinoco, I., Jr. (1992). Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops, *Nucleic Acids Res.* **20**, 819–824.

Applequist, J. (1963). On the Helix-Coil Equilibrium in Polypeptides, *J. Chem. Phys.* **38**, 934–941.

Arnott, S., Chandrasekaran, R., Hall, I. H., Puigjaner, L. C., Walker, J. K. and Wang, M. (1982). DNA Secondary Structures: Helices, Wrinkles, and Junctions, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 53–65.

Banerjee, A. R., Jaeger, J. A., and Turner, D. H. (1993). Thermal unfolding of a group I ribozyme: The low-temperature transition is primarily disruption of tertiary structure, *Biochemistry* **32**, 153–163.

Battiste, J. L., Tan, R., Frankel, A. D., and Williamson, J. R. (1994). Binding of an HIV Rev peptide to Rev responsive element RNA induces formation of purine-purine base pairs, *Biochemistry* **33**, 2741–2747.

Behe, M. and Felsenfeld, G. (1981). Effects of methylation on a synthetic polynucleotide: The B–Z transition in poly(dG-m5dC):poly(dG-m5dC), *Proc. Natl. Acad. Sci. USA* **78**, 1619–1623.

Berman, H. M. (1981). Conformational Principles of Nucleic Acids, in *Topics in Nucleic Acid Structure*, Part 1, Chapter 1 Neidle, S., Ed., Macmillan, London, UK.

Bernasconi, C. F. (1976). *Relaxation Kinetics*, Academic Press, New York.

Bevilacqua, P. C., Johnson, K. A., and Turner, D. H. (1993). Cooperative and Anticooperative Binding to a Ribozyme, *Proc. Natl. Acad. Sci. USA* **90**, 8357–8361.

Bevilacqua, P. C., Kierzek, R., Johnson, K. A., and Turner, D. H. (1992). Dynamics of Ribozyme Binding of Substrate Revealed by Fluorescence-Detected Stopped-Flow Methods, *Science* **258**, 1355–1358.

Bevilacqua, P. C., Li, Y., and Turner, D. H. (1994). Fluorescence-Detected Stopped Flow with a Pyrene Labeled Substrate Reveals that Guanosine Facilitates Docking of the 5′ Cleavage Site into a High Free Energy Binding Mode in the *Tetrahymena* Ribozyme, *Biochemistry* **33**, 11340–11348.

Bevilacqua, P. C. and Turner, D. H. (1991). Comparison of Binding of Mixed Ribose-Deoxyribose Analogues of CUCU to a Ribozyme and to GGAGAA by Equilibrium Dialysis: Evidence for Ribozyme Specific Interactions with 2′ OH Groups, *Biochemistry* **30**, 10632–10640.

Blagoi, Y. P., Sorokin, V. A., Valeyev, V. A., Khomenko, S. A., and Gladchenko, G. O. (1978). Magnesium Ion Effect on the Helix–Coil Transition of DNA, *Biopolymers* **17**, 1103–1118.

Blake, R. D. (1996). Denaturation of DNA, in *Encyclopedia of Molecular Biology and Molecular Medicine*, Vol. 2, Meyers, R. A., Ed. VCH, New York, 1–19.

Blake, R. D. and Delcourt, S. G. (1987). Loop Energy in DNA, *Biopolymers* **26**, 2009– 2026.

Blake, R. D. and Haydock, P. V. (1979). Effect of Sodium Ion on the High-Resolution Melting of Lambda DNA, *Biopolymers* **18**, 3089–3109.

Bloomers, M. J. J., Walters, J. A. L. I., Haasnoot, C. A. G., Aelen, J. M. A., van der Marel, G. A., van Boom, J. J., and Hilbers, C. W. (1989). Effects of Base Sequence on the Loop Folding in DNA Hairpins, *Biochemistry* **28**, 7491–7498.

Bloomfield, V. A., Crothers, D. M. and Tinoco, I., Jr. (1974). *Physical Chemistry of Nucleic Acids*, Chapter 6, Harper and Row, New York.

Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974). Stability of Ribonucleic Acid Double-Stranded Helices, *J. Mol. Biol.* **86**, 843–853.

Breslauer, K. J. and Bina-Stein, M. (1977). Relaxation Kinetics of the Helix-Coil Transition of a Self-Complementary Ribo-Oligonucleotide: A_2U_2, *Biophysical Chem.* **7**, 211–216.

Breslauer, K. J., Frank, R., Blöcker, H., and Marky, L. A. (1986). Predicting DNA Duplex Stability from the Base Sequence, *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.

Breslauer, K. J., Freire, E., and Straume, M. (1992). Calorimetry: A Tool for DNA and Ligand-DNA Studies, *Methods Enzymol.* **211**, 533–567.

Broitman, S. L., Im, D. D., and Fresco, J. R. (1987). Formation of the triple-stranded polynucleotide helix, poly (A:A:U), *Proc. Natl. Acad. Sci. USA* **84**, 5120–5124.

Burkard, M. E., Kierzek, R., and Turner, D. H. (1999). Thermodynamics of Unpaired Terminal Nucleotides on Short RNA Helixes Correlates with Stacking at Helix Termini in Larger RNAs, *J. Mol. Biol.* **290**, 967-982.

Calladine, C. R. (1982). Mechanics of Sequence-Dependent Stacking of Bases in B-DNA, *J. Mol. Biol.* **161**, 343–352.

Cantor, C. R. and Schimmel, P. R. (1980). *Biophysical Chemistry*, Part III, Chapter 23, Freeman, San Francisco, CA.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996). Crystal Structure of a Group I Ribozyme Domain: Principles of RNA Packing, *Science* **273**, 1678–1685.

Cate, J. H., Hanna, R. L., and Doudna, J. A. (1997). A Magnesium Ion Core at the Heart of a Ribozyme Domain, *Nature Struct. Biol.* **4**, 553–558.

Cech, T. R. (1990). Self-splicing of Group I Introns, *Annu. Rev. Biochem.* **59**, 543–568.

Cech, T. R. and Bass, B. L. (1986). Biological Catalysis By RNA, *Annu. Rev. Biochem.* **55**, 599–629.

Cech, T. R., Damberger, S. H., and Gutell, R. R. (1994). Representation of the Secondary and Tertiary Structure of Group I Introns, *Nature Struct. Biol.* **1**, 273–280.

Chaires, J. B. and Sturtevant, J. M. (1986). Thermodynamics of the B to Z transition in poly (m5dG-dC), *Proc. Natl. Acad. Sci. USA* **83**, 5479–5483.

Chastain, C. M. and Tinoco, I., Jr. (1991). Structural elements in RNA, *Prog. Nucleic Acid Res. Mol. Biol.* **41**, 131–177.

Chastain, C. M. and Tinoco, I., Jr. (1992). Poly(rA) Binds Poly(rG)·Poly(rC) to Form a Triple Helix, *Nucleic Acids Res.* **20**, 315–318.

Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X. C., Stern, D., Winkler, J., Lockhart, D. J., Morris, M. S., and Fodor, S. P. A. (1996). Accessing Genetic Information with High-Density DNA Arrays, *Science* **274**, 610–614.

Cheong, C., Tinoco, I., Jr., and Chollet, A. (1988). Thermodynamic Studies of Base Pairing Involving 2,6-Diaminopurine, *Nucleic Acids Res.* **16**, 5115–5122.

Christophe, D., Cabrer, B., Bacolla, A., Targovnik, H., Pohl, V., and Vassart, G. (1985). An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream from the human thyroglobulin gene, *Nucleic Acids Res.* **13**, 5127–5144.

Chu, Y. G. and Tinoco, I., Jr. (1983). Temperature-Jump Kinetics of the dC-G-T-G-A-A-T-T-C-G-C-G Double Helix Containing a G:T Base Pair and the dC-G-C-A-G-A-A-T-T-C-G-C-G Double Helix Containing an Extra Adenine, *Biopolymers* **22**, 1235-1246.

Clore, G. M., Gronenborn, A. M., Piper, E. A., McLaughlin, L. W., Graeser, E., and Van Boom, J. H. (1984). The Solution Structure of a RNA Pentadecamer Comprising the Anticodon Loop and Stem of Yeast $tRNA^{Phe}$, *Biochem. J.* **221**, 737–751.

Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J., and Hogan, M. E. (1988). Site-Specific Oligonucleotide Binding Represses Transcription of the Human c-myc Gene in Vitro, *Science* **241**, 456–459.

Costa, M. and Michel, F. (1995). Frequent Use of the Same Tertiary Motif by Self-Folding RNAs, *EMBO J.* **14**, 1276–1285.

Coutts, S. M., Gangloff, J., Dirheimer, G. (1974). Conformational Transitions in $tRNA^{Asp}$ (Brewer's Yeast). Thermodynamic, Kinetic, and Enzymatic Measurements on Oligonucleotide Fragments and the Intact Molecule, *Biochemistry* **13**, 3938–3948.

Coutts, S. M., (1971). Thermodynamics and Kinetics of G:C Base Pairing in the Isolated Extra Arm of Serine-Specific Transfer RNA from Yeast, *Biochem. Biophys. Acta* **232**, 94–106.

Craig, M. E., Crothers, D. M., and Doty, P. (1971). Relaxation Kinetics of Dimer Formation by Self Complementary Oligonucleotides, *J. Mol. Biol.* **62**, 383–401.

Crick, F. H. C. (1966). Codon–Anticodon Pairing: The Wobble Hypothesis, *J. Mol. Biol.* **19**, 548–555.

Crothers, D. M. and Cole, P. E. (1978). In *Transfer RNA*, Altman, S., Ed., MIT Press, Cambridge, MA, pp. 196–247.

Crothers, D. M., Cole, P. E., Hilbers, C. W., and Shulman, R. G. (1974). The Molecular Mechanism of Thermal Unfolding of *Escherichia coli* Formylmethionine Transfer RNA, *J. Mol. Biol.* **87**, 63–88.

Cruz, P., Hall, K., Puglisi, J., Davis, P., Hardin, C. C., Trulson, M. O., Mathies, R. A., Tinoco, I., Jr., Johnson, W. C., Jr., and Neilson, T. (1986). The Left-Handed Z-form of Double Stranded RNA, in *Biomolecular Stereodynamics IV*, Sarma, R. H. and Sarma, M. H., Eds., Adenine Press, Albany, NY, pp. 179–200.

Davis, R. C. and Tinoco, I., Jr. (1968). Temperature-dependent Properties of Dinucleoside Phosphates, *Biopolymers* **6**, 223–242.

Delcourt, S. G. and Blake, R. D. (1991). Stacking Energies in DNA, *J. Biol. Chem.* **266**, 15160–15169.

Dewey, T. G., Raymond, D. A., and Turner, D. H. (1979). Laser Temperature Jump Study of Solvent Effects on Proflavin Stacking, *J. Am. Chem. Soc.* **101**, 5822– 5826.

Dewey, T. G. and Turner, D. H. (1979). Laser Temperature-Jump Study of Stacking in Adenylic Acid Polymers, *Biochemistry* **18**, 5757–5762.

Dickerson, R. E. (1983). Base Sequence and Helix Structure Variation in B and A DNA, *J. Mol. Biol.* **166**, 419–441.

Dickerson, R. E., Drew, H. R., Conner, B. N., Kopka, M. L., and Pjura, P. E. (1982). Helix Geometry and Hydration in A-DNA, B-DNA, and Z-DNA, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 13–24.

Dieckmann, T., Suzuki, E., Nakamura, G. K., and Feigon, J. (1996). Solution Structure of an ATP-binding RNA Aptamer Reveals a Novel Fold, *RNA*, **2**, 628–640.

Doktycz, M. J., Goldstein, R. F., Paner, T. M., Gallo, F. J., and Benight, A. S. (1992). Studies of DNA Dumbbells. I. Melting Curves of 17 DNA Dumbbells with Different Duplex Stem Sequences Linked by T4 Endloops: Evaluation of the Nearest-Neighbor Stacking Interactions in DNA, *Biopolymers* **32**, 849–864.

Edwards, E. L., Ratliff, R. L., and Gray, D. M. (1988). Circular Dichroism Spectra of DNA Oligomers Show that Short Interior Stretches of C:C$^+$ Base Pairs Do Not Form in Duplexes with A:T Base Pairs, *Biochemistry* **27**, 5166–5174.

Eigen, M. and De Maeyer, L. (1963). in *Techniques of Organic Chemistry*, Vol. VIII, Part 2, Friess, S. L., Lewis, E. S. and Weissberger, A., Eds., Wiley-Interscience, New York, p. 895.

Erie, D. A., Jones, R. A., Olson, W. K., Sinha, N. K., and Breslauer, K. J. (1989). Melting Behavior of a Covalently Closed, Single-Stranded, Circular DNA, *Biochemistry* **28** , 268–273.

Erie, D., Sinha, N., Olson, W., Jones, R., and Breslauer, K. J. (1987). A Dumbbell-Shaped, Double-Hairpin Structure of DNA: A Thermodynamic Investigation, *Biochemistry* **26**, 7150–7159.

Fan, P., Suri, A. K., Fiala, R., Live, D., and Patel, D. J. (1996). Molecular Recognition in the FMN–RNA Aptamer Complex, *J. Mol. Biol.* **258**, 480–500.

Felsenfeld, G., Davies, D. R., and Rich, A. (1957). Formation of a Three-Stranded Polynucleotide Molecule, *J. Am. Chem. Soc.* **79**, 2023–2024.

Felsenfeld, G. and Todd Miles, H. (1967). The Physical and Chemical Properties of Nucleic Acids, *Annu. Rev. Biochem.* **36**, 407–448.

Filimonov, V. V. and Privalov, P. L. (1978). Thermodynamics of Base Interaction in $(A)_n$ and $(A:U)_n$, *J. Mol. Biol.* **122**, 465–470.

Fink, T.R. and Crothers, D. M. (1972). Free Energy of Imperfect Nucleic Acid Helices. I. The Bulge Defect, *J. Mol. Biol.* **66**, 1–12.

Fisher, M. E. (1966). Effect of Excluded Volume on Phase Transitions in Biopolymers, *J. Chem. Phys.* **45**, 1469–1473.

Flor, P., Flanegan, J. B., and Cech, T. R. (1989). A Conserved Base-pair within Helix P4 of the Tetrahymena Ribozyme Helps to Form the Tertiary Structure Required for Self-Splicing. *EMBO J.* **8**, 3391–3399.

Frank-Kamenetskii, M. D. (1971). Simplification of the Empirical Relationship between Melting Temperatures of DNA, Its GC Content and Concentration of Sodium Ions in Solution, *Biopolymers* **10**, 2623–2624.

Freier, S. M., Albergo, D. D., and Turner, D. H. (1983a). Solvent Effects on the Dynamics of $(dG\text{-}dC)_4$, *Biopolymers* **22**, 1107–1131.

Freier, S. M., Alkema, D., Sinclair, A., Neilson, T., and Turner, D. H. (1985). Contributions of Dangling End Stacking and Terminal Base-Pair Formation to the Stabilities of XGGCCp, XCCGGp, XGGCCYp, and XCCGGYp Helixes, *Biochemistry* **24**, 4533–4539.

Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., and Turner, D. H. (1983). Effects of 3′ Dangling End Stacking on the Stability of GGCC and CCGG Double Helices, *Biochemistry* **22**, 6198–6206.

Freier, S. M., Hill, K. O., Dewey, T. G., Marky, L. A., Breslauer, K. J., and Turner, D. H. (1981). Solvent Effects on the Kinetics and Thermodynamics of Stacking in Polylcytidylic Acid, *Biochemistry* **20**, 1419–1426.

Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986). Free Energy Contributions of G:U and Other Terminal Mismatches to Helix Stability, *Biochemistry* **25**, 3209–3213.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986b). Improved Free-Energy Parameters for Predictions of RNA Duplex Stability, *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.

Freier, S. M., Sinclair, A., Neilson, T., and Turner, D. H. (1985b). Improved Free Energies for G:C Base-Pairs, *J. Mol. Biol.* **185**, 645–647.

Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986c). Stability of XGCGCp, GCGCYp, and XGCGCYp Helixes: An Empirical Estimate of the Energetics of Hydrogen Bonds in Nucleic Acids, *Biochemistry* **25**, 3214–3219.

Gaffney, B. L. and Jones, R. A. (1989). Thermodynamic Comparison of the Base Pairs Formed by the Carcinogenic Lesion O6-Methylguanine with Reference both to Watson–Crick Pairs and to Mismatched Pairs, *Biochemistry* **28**, 5881–5889.

Gaffney, B. L., Marky, L. A., and Jones, R. A. (1984). The Influence of the Purine 2-Amino Group on DNA Conformation and Stability–II. Synthesis and Physical Characterization of d[CGT(2-NH_2)ACG], d[CGU(2-NH_2)ACG], and d[CGT(2-NH_2)AT(2-NH_2)ACG], *Tetrahedron* **40**, 3–13.

Gellert, M., Lipsett, M. N., and Davies, D. R. (1962). Helix Formation by Guanylic Acid, *Proc. Natl. Acad. Sci. USA* **48**, 2013–2018.

Gellert, M., Mizuuchi, K., O'Dea, M. H., Ohmori, H., and Tomizawa, J. (1979). DNA Gyrase and DNA Supercoiling, *Cold Spring Harbor Symp. Quant. Biol.* **43**, 35–40.

Goldstein, R. F. and Benight, A. S. (1992). How many Numbers are Required to Specify Sequence-Dependent Properties of Polynucleotides? *Biopolymers* **32**, 1679–1693.

Gotoh, O. (1983). Prediction of Melting Profiles and Local Helix Stability for Sequenced DNA, *Adv. Biophys.* **16**, 1–52.

Gouy, M. (1986). Secondary Stucture Prediction of RNA, in *Nucleic Acid and Protein Sequence Analysis: A Practical Approach*, Bishop, M. J. and Rawlings, C. J., Eds., IRL Press, Oxford, UK, pp. 259–284.

Gralla, J. and Crothers, D. M. (1973a). Free Energy of Imperfect Nucleic Acid Helices II. Small Hairpin Loops, *J. Mol. Biol.* **73**, 497–511.

Gralla, J. and Crothers, D. M. (1973b). Free Energy of Imperfect Nucleic Acid Helices, *J. Mol. Biol.* **78**, 301–319.

Gray, D. M. (1997a). Derivation of Nearest-Neighbor Properties from Data on Nucleic Acid Oligomers. I. Simple Sets of Independent Sequences and the Influence of Absent Nearest Neighbors, *Biopolymers* **42**, 783–793.

Gray, D. M. (1997b). Derivation of Nearest-Neighbor Properties from Data on Nucleic Acid Oligomers. II. Thermodynamic Parameters of DNA-RNA Hybrids and DNA Duplexes, *Biopolymers* **42**, 795–810.

Gray, D. M., Cui, T., and Ratliff, R. L. (1984). Circular dichroism measurements show that C:C$^+$ base pairs can coexist with A:T base pairs between antiparallel strands of an oligodeoxynucleotide double-helix, *Nucleic Acids Res.* **12**, 7565–7580.

Gray, D. M., Vaughn, M., Ratliff, R. L., and Hayes, F. N. (1980). Circular dichorism spectra show that repeating dinucleotide DNAs may form helices in which every other base is looped out, *Nucleic Acids Res.* **8**, 3695–3707.

Griffin, L. C. and Dervan, P. B. (1989). Recognition of Thymine·Adenine Base Pairs by Guanine in a Pyrimidine Triple Helix Motif, *Science* **245**, 967–971.

Groebe, D. R. and Uhlenbeck, O. C. (1988). Characterization of RNA Hairpin Loop Stability, *Nucleic Acids Res.* **16**, 11725–11735.

Groebe, D. R. and Uhlenbeck, O. C. (1989). Thermal Stability of RNA Hairpins Containing a Four-Membered Loop and a Bulge Nucleotide, *Biochemistry* **28**, 742–747.

Gultyaev, A. P., van Batenburg, F. H. D., and Pleij, C. W. A. (1995). The Computer Simulation of RNA Folding Pathways using a Genetic Algorithm *J. Mol. Biol.* **250**, 37–51.

Guschlbauer, W. (1975). Protonated polynucleotide structures. Thermodynamics of the melting of the acid form of polycytidylic acid, *Nucleic Acids Res.* **2**, 353–360.

Gutell, R. R., Larsen, N., and Woese, C. R. (1994). Lessons from an Evolving RNA: 16S and 23S rRNA Structures from a Comparative Perspective, *Microbiol. Rev.* **58**, 10–26.

Hall, K. B. and Maestre, M. F. (1984). Temperature-Dependent Reversible Transition of Poly(dCdG): Poly(dCdG) in Ethanolic and Methanolic Solutions, *Biopolymers* **23**, 2127–2139.

Hamaguchi, K. and Geiduschek, E. P. (1962). The Effect of Electrolytes on the Stability of the Deoxyribonucleate Helix, *J. Am. Chem. Soc.* **84**, 1329–1338.

Hardin, C. C., Henderson, E., Watson, T., and Prosser, J. R. (1991). Monovalent Cation Induced Structural Transitions in Telomeric DNAs: G-DNA Folding Intermediates, *Biochemistry* **30**, 4460–4472.

Hare, D., Shapiro, L. and Patel, D. J. (1986). Extrahelical Adenosine Stacks into Right-Handed DNA: Solution Conformation of the d(C-G-C-A-G-A-G-C-T-C-G-C-G) Duplex Deduced from Distance Geometry Analysis of Nuclear Overhauser Effect Spectra, *Biochemistry* **25**, 7456–7464.

Hartman, K. A., Jr. and Rich, A. (1965). The Tautomeric Form of Helical Polyribocytidylic Acid, *J. Am. Chem. Soc.* **87**, 2033–2039.

He, L., Kierzek, R., SantaLucia, J., Jr., Walter, A. E., and Turner, D. H. (1991). Nearest-Neighbor Parameters of G.U Mismatches: $\overset{5'GU3'}{3'UG5'}$ Is Destabilizing in the Contexts $\overset{CGUG}{GUGC}$, $\overset{UGUA}{AUGU}$, and $\overset{AGUU}{UUGA}$, but Stabilizing in $\overset{GGUC}{CUGG}$, *Biochemistry* **30**, 11124–11132.

Herschlag, D. and Cech, T. R. (1990). Catalysis of RNA Cleavage by the Tetrahymena thermophila Ribozyme. A. Kinetic Description of the Reaction of an RNA Substrate Complementary to the Active Site, *Biochemistry* **29**, 10159–10171.

Herskovits, T. T. and Bowen, J. J. (1974). Solution Studies of the Nucleic Acid Bases and Related Model Compounds. Solubility in Aqueous Urea and Amide Solutions, *Biochemistry* **13**, 5474–5483.

Herskovits, T. T. and Harrington, J. P. (1972). Solution Studies of the Nucleic Acid Bases and Related Model Compounds. Solubility in Aqueous Alcohol and Glycol Solutions, *Biochemistry* **11**, 4800–4811.

Heus, H. A. and Pardi, A. (1991). Structural Features that give Rise to the Unusual Stability of RNA Hairpins Containing GNRA Loops, *Science* **253**, 191–194.

Hickey, D. R. and Turner, D. H. (1985a). Effects of Terminal Mismatches on RNA Stability: Thermodynamics of Duplex Formation for ACCGGGp, ACCGGAp, and ACCGGCp, *Biochemistry* **24**, 3987–3991.

Hickey, D. R. and Turner, D. H. (1985b). Solvent Effects on the Stability of A_7U_7p, *Biochemistry* **24**, 2086–2094.

Hilbers, C. W., Haasnoot, C. A. G., de Bruin, S. H., Joordens, J. J. M., Van Der Marel, G. A., and Van Boom, H. H. (1985). Hairpin Formation in Synthetic Oligonucleotides, *Biochimie* **67**, 685–695.

Hilbers, C. W., Heus, H. A., van Dongen, M. J. P., and Wijmenga, S. S. (1994). The Hairpin Elements of Nucleic Acid Structure: DNA and RNA Folding, *Nucleic Acids Mol. Biol.* **8**, 56–104.

Hoggett, J. G. and Maass, G. (1971). Thermodynamics and Kinetics of the Formation of a Hybrid Double-Helix of Oligomers of Riboadenylic and Deoxyribothymidylic Acids of Definite Chain Lengths, *Ber. Bunsengers Phys. Chem.* **75**, 45–54.

Holcomb, D. N. and Timascheff, S. N. (1968). Temperature Dependence of the Hydrogen Ion Equilibria in Poly(riboadenylic Acid), *Biopolymers* **6**, 513–529.

Htun, H. and Dahlberg, J. E. (1988). Single Strands, Triple Strands, and Kinks in H-DNA, *Science* **241**, 1791–1796.

Hud, N. V., Smith, F. W., Aret, F. A. L., and Feigon, J. (1996). The Selectivity for K^+ versus Na^+ in DNA Quadruplexes is Dominated by Relative Free Energies of Hydration: A Thermodynamic Analysis by ^1H NMR, *Biochemistry* **35**, 15383–15390.

Inman, R. B. (1964). Transitions of DNA Homopolymers, *J. Mol. Biol.* **9**, 624–637.

Inners, L. D. and Felsenfeld, G. (1970). Conformation of Polyribouridylic Acid in Solution, *J. Mol. Biol.* **50**, 373–389.

Ivanov, V. I., Minchenkova, L. E., Minyat, E. E., Frank- Kamenetskii, M. D., and Schyolkina, A. K. (1974). The B to A Transition of DNA in Solution, *J. Mol. Biol.* **87**, 817–833.

Jacobson, H. and Stockmayer, W. H. (1950). Intramolecular Reaction in Polycondensations, I. The Theory of Linear Systems, *J. Chem. Phys.* **18**, 1600–1606.

Jaeger, J. A., Turner, D. H., and Zuker, M. (1989). Improved Predictions of Secondary Structures for RNA, *Proc. Natl. Acad. Sci. USA* **86**, 7706–7710.

Jaeger, J. A., Zuker, M., and Turner, D. H. (1990). Melting and Chemical Modification of a Cyclized Self-Splicing Group I Intron: Similarities of Structures in 1 *M* Na^+, in 10 m*M* Mg^{2+}, and in the Presence of Substrate, *Biochemistry*, **29**, 10148–10158.

Jaeger, J. A., Westhof, E. and Michel, F. (1993). Monitoring of the cooperative unfolding of the sunY group I intron of bacteriophage T4, *J. Mol. Biol.* **234**, 331–346.

Jencks, W. P. (1975). Binding Energy, Specificity, and Enzymic Catalysis: The Circe Effect, *Adv. Enzymol.* **43**, 219–410.

Jiang, F., Kumar, R. A., Jones, R. A., and Patel, D. J. (1996). Structural Basis of RNA Folding and Recognition in an AMP–RNA Aptamer Complex, *Nature (London)* **382**, 183–86.

Jin, R., Chapman, W. H., Jr., Srinivasan, A. R., Olson, W. K., Breslow, R., and Breslauer, K. J. (1993). Comparative Spectroscopic, Calorimetric, and Computational Studies of Nucleic Acid Complexes with 2',5" versus 3',5"-Phosphodiester Linkages, *Proc. Natl. Acad. Sci. USA* **90**, 10568–10572.

Jin, R., Gaffney, B. L., Wang, C., Jones, R. A., and Breslauer, K. J. (1992). Thermodynamics and Structure of a DNA Tetraplex: A Spectroscopic and Calorimetric Study of the Tetramolecular Complexes of $d(TG_3T)$ and $d(TG_3T_2G_3T)$, *Proc. Natl. Acad. Sci. USA* **89**, 8832–8836.

Johnston, B. H. (1988). The S1-Sensitive Form of $d(C\text{-}T)_n{\cdot}d(A\text{-}G)_n$: Chemical Evidence for a Three-Stranded Structure in Plasmids, *Science* **241**, 1800–1804.

Jovin, T. M. and Soumpasis, D. M. (1987). The Transition Between B-DNA and Z-DNA, *Annu. Rev. Phys. Chem.* **38**, 521–560.

Jucker, F. M. and Pardi, A. (1995). Solution Structure of the CUUG Hairpin Loop: A Novel RNA Tetraloop Motif, *Biochemistry* **34**, 14416–14427.

Kao, T. H. and Crothers, D. M. (1980). A proton-coupled conformational switch of Escherichia coli 5S ribosomal RNA, *Proc. Natl. Acad. Sci. USA* **77**, 3360–3364.

Kierzek, R., Caruthers, M. H., Longfellow, C. E., Swinton, D., Turner, D. H., and Freier, S. M. (1986). Polymer-Supported RNA Synthesis and Its Application to Test the Nearest-Neighbor Model for Duplex Stability, *Biochemistry* **25**, 7840–7846.

Kierzek, R., He, L., and Turner, D. H. (1992). Association of 2'-5' Oligoribonucleotides, *Nucleic Acids Res.* **29**, 1685–1690.

Kim, S. H. and Cech, T. R. (1987). Three-dimensional Model of the Active Site of the Self-Splicing rRNA Precursor of Tetrahymena, *Proc. Natl. Acad. Sci. USA* **84**, 8788–8792.

Kim, J., Cheong, C., and Moore, P. B. (1991). Tetramerization of an RNA Oligonucleotide Containing a GGGG Sequence, *Nature (London)* **351**, 331-332.

Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H.-J., Seeman, N. C., and Rich, A. (1974). Three-Dimensional Tertiary Structure of Yeast Phenylalanine Transfer RNA, *Science* **185**, 435–440.

Klump, H. H. (1990). Saenger, W. Ed., Landolt-Börnstein, New Series, Nucleic Acids, Subvol. C, Spectroscopic and Kinetic Data, in VII Biophysics, Vol 1, Springer-Verlag, Berlin, Germany, pp. 241–256.

Klump, H. and Burkart, W. (1977). Calorimetric Measurements of the Transition Enthalpy of DNA in Aqueous Urea Solutions, *Biochim. Biophys. Acta* **475**, 601–604.

Klump, H. H. and Jovin, T. M. (1987). Formation of a Left-Handed RNA Double Helix: Energetics of the A-Z Transition of Poly[r(G-C)] in Concentrated $NaClO_4$ Solutions, *Biochemistry* **26**, 5186–5190.

Knitt, D. S., Narlikar, G. J., and Herschlag, D. (1994). Dissection of the Role of the Conserved G•U Pair in Group I RNA Self-Splicing, *Biochemistry* **33**, 13864–13879.

Kondo, N. S. and Danyluk, S. S. (1976). Conformational Properties of Adenylyl- 3'5'-adenosine in Aqueous Solution, *Biochemistry* **15**, 756–768.

Krakauer, H. (1974). A Thermodynamic Analysis of the Influence of Simple Mono- and Divalent Cations on the Conformational Transitions of Polynucleotide Complexes, *Biochemistry* **13**, 2579–2589.

Krakauer, H. and Sturtevant, J. M. (1968). Heats of the Helix–Coil Transitions of the Poly A-Poly U Complexes, *Biopolymers* **6**, 491–512.

Laing, L. G. and Draper, D. E. (1994). Thermodynamics of RNA Folding in a Conserved Ribosomal RNA Domain, *J. Mol. Biol.* **237**,560–576.

Lee, J. S., Johnson, D. A., and Morgan, A. R. (1979). Complexes formed by $(pyrimidine)_n{\cdot}(purine)_n$ DNAs on lowering the pH are three-stranded, *Nucleic Acids Res.* **6**, 3073–3091.

Lee, C. H. and Tinoco, I., Jr. (1980). Conformation Studies of 13 Trinucleoside Diphosphates by 360 MHz PMR Spectroscopy. A Bulged Base Conformation. I. Base Protons and H1' Protons, *Biophys. Chem.* **11**, 283–294.

Legault, P. and Pardi, A. (1994). In-Situ Probing of Adenine Protonation in RNA by C-13 NMR, *J. Am. Chem. Soc.* **116**, 8390-8391.

Leontis, N. B., Kwok, W. and Newman, J. S. (1991). Stability and Structure of Three-way DNA Junctions Containing Unpaired Nucleotides, *Nucleic Acids Res.* **19** , 759–766.

Lerman, L. S., Fischer, S. G., Hurley, I., Silverstein, K., and Lumelsky, N. (1984). Sequence-Determined DNA Separations, *Annu. Rev. Biophys. Bioeng.* **13**, 399–423.

Levine, L., Gordon, J. A., and Jencks, W. P. (1963). The Relationship of Structure to the Effectiveness of Denaturing Agents for Deoxyribonucleic Acid, *Biochemistry* **2**, 168–175.

Lipsett, M. N. (1963). The Interactions of Poly C and Guanine Trinucleotide, *Biochem. Biophys. Res. Commun.* **11**, 224–228.

Lipsett, M. N. (1964). Complex Formation between Polycytidylic Acid and Guanine Oligonucleotides, *J. Biol. Chem.* **239**, 1256–1260.

Longfellow, C. E., Kierzek, R. and Turner, D. H. (1990). Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides, *Biochemistry* **29**, 278–285.

Lu, M. and Draper, D. E. (1994). Bases defining an ammonium and magnesium ion dependent tertiary structure within the large subunit ribosomal RNA, *J. Mol. Biol.* **244**, 572–585.

Lu, M., Guo, Q. and Kallenbach, N. R. (1993). Thermodynamics of G-Tetraplex Formation by Telomeric DNAs, *Biochemistry* **32**, 598–601.

Lu, M., Guo, Q., Marky, C. A., Seeman, N. C., and Kallenbach, N. R. (1992). Thermodynamics of DNA Branching, *J. Mol. Biol.* **223** , 781–789.

Lück, R., Steger, G., and Riesner, D. (1996). Thermodynamic Prediction of Conserved Secondary Structure: Application to the RRE Element of HIV, the tRNA-like Element of CMV and the mRNA of Prion Protein, *J. Mol. Biol.* **258**, 813–826.

Manning, G. S. (1976). On the application of polyelectrolyte limiting laws to the helix–coil transition of DNA. V. Ionic effects on renaturation kinetics, *Biopolymers* **15**, 1333–1343.

Manning, G. S. (1978). The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides, *Q. Rev. Biophys.* **11**, 179–246.

Marky, L. A. and Breslauer, K. J. (1987). Calculating Thermodynamic Data for Transitions of any Molecularity from Equilibrium Melting Curves, *Biopolymers* **26**, 1601–1620.

Marky, L. A., Kallenbach, N. R., McDonough, K. A., Seeman, N. C., and Breslauer, K. J. (1987). The Melting Behavior of a DNA Junction Structure: A Calorimetric and Spectroscopic Study, *Biopolymers* **26**, 1621–1634.

Marmur, J. and Doty, P. (1962). Determination of the Base Composition of Deoxyribonucleic Acid from its Thermal Melting Temperature, *J. Mol. Biol.* **5**, 109–118.

Martin, F. H., Castro, M. M., Aboul-ela, F., and Tinoco, I., Jr. (1985). Base Pairing Involving Deoxyinosine: Implications for Probe Design, *Nucleic Acids Res.* **13**, 8927–8938.

Martin, F. H. and Tinoco, I., Jr. (1980). DNA–RNA Hybrid Duplexes Containing Oligo (dA:rU) Sequences Are Exceptionally Unstable and May Facilitate Termination of Transcription, *Nucleic Acids Res.* **8**, 2295–2299.

Martin, F. H., Uhlenbeck, O. C., and Doty, P. (1971). Self-Complementary Oligoribonucleotides: Adenylic Acid-Uridylic Acid Block Copolymers, *J. Mol. Biol.* **57**, 201–215.

Mathews, D. H., Banerjee, A. R., Luan, D. D., Eickbush, T. H., and Turner, D. H. (1997). Secondary Structure Model of the RNA Recognized by the Reverse Transcriptase from the R2 Retrotransposable Element, *RNA* **3**, 1–16.

Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999). Expanded Sequence Dependence of Thermodynamic Parameters Improves Predictions of RNA Secondary Structure, *J. Mol. Biol.* **288**, 911–940.

McConnell, T. S., Cech, T. R., and Herschlag, D. (1993). Guanosine Binding to the *Tetrahymena* Ribozyme: Thermodynamic Coupling with Oligonucleotide Binding, *Proc. Natl. Acad. Sci. USA* **90**, 8362–8366.

Mellema, R. J., van der Woerd, R., van der Marel, G. A., van Boom, J. H., and Altona, C. (1984). Proton NMR Study and Conformational Analysis of d(CGT), d(TCG), and d(CGTCG) in Aqueous Solution. The Effect of a Dangling Thymidine and of a Thymidine Mismatch on DNA Mini-duplexes, *Nucleic Acids Res.* **12**, 5061–5078.

Michel, F., Ellington, A. D., Couture, S., and Szostak, J. W. (1990). Phylogenetic and Genetic Evidence for Base-Triples in the Catalytic Domain of Group I Introns, *Nature (London)* **347**, 578–580.

Michel, F. and Westhof, E. (1990). Modelling of the Three-dimensional Architecture of Group I Catalytic Introns Based on Comparative Sequence Analysis, *J. Mol. Biol.* **216**, 585–610.

Michelson, A. M., Massoulie, J., and Guschlbauer, W. (1967). Synthetic Polynucleotides, *Prog. Nucleic Acids Res. Mol. Biol.* **6**, 83-141.

Miller, J. H. and Sobell, H. M. (1966). A Molecular Model for Gene Repression, *Proc. Natl. Acad. Sci. USA* **55**, 1201–1205.

Mirkin, S. M. and Frank-Kamenetskii, M. D. (1994). H-DNA and Related Structures, *Annu. Rev. Biophys. Biomolec. Struct.* **23**, 541–576.

Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., and Frank-Kamenetskii, M. D. (1987). DNA H form requires a homopurine–homopyrimidine mirror repeat, *Nature (London)* **330**, 495–497.

Moazed, D., Stern, S., Noller, H. F. (1986). Rapid Chemical Probing of Conformation in 16S Ribosomal RNA and 30S Ribosomal Subunits Using Primer Extension, *J. Mol. Biol.* **187**, 399–416.

Moran, S., Ren, R. X.-F., and Kool, E. T. (1997). A Thymidine Triphosphate Shape Analog Lacking Watson–Crick Pairing Ability is Replicated with High Sequence Selectivity, *Proc. Natl. Acad. Sci. USA* **94**, 10506–10511.

Moran, S., Kierzek, R., and Turner, D. H. (1993). Binding of Guanosine and 3' Splice Site Analogues to a Group I Ribozyme: Interactions with Functional Groups of Guanosine and with Additional Nucleotides, *Biochemistry* **32**, 5247–5256.

Morden, K. M., Chu, Y. G., Martin, F. H. and Tinoco, I., Jr. (1983). Unpaired Cytosine in the Deoxyoligonucleotide Duplex dCA_1CA_1G:dCT_4G is Outside of the Helix, *Biochemistry* **22**, 5557–5563.

Murphy, F. L. and Cech, T. R. (1994). GAAA Tetraloop and Conserved Bulge Stabilize Tertiary Structure of a Group I Intron Domain, *J. Mol. Biol.* **236**, 49–63.

Murphy, F. L. and Cech, T. R. (1993). An Independently Folding Domain of RNA Tertiary Structure within the Tetrahymena Ribozyme, *Biochemistry* **32**, 5291–5300.

Narlikar, G. J., Gopalakrishnan, V., McConnell, T. S., Usman, N., and Herschlag, D. (1995). Use of binding energy by an RNA enzyme for catalysis by positioning and substrate destabilization, *Proc. Natl. Acad. Sci. USA* **92**, 3668–3672.

Nelson, J. W., Martin, F. H. and Tinoco, I., Jr. (1981). DNA and RNA Oligomer Thermodynamics: The Effect of Mismatched Bases on Double-Helix Stability, *Biopolymers* **20**, 2509–2531.

Nelson, J. W. and Tinoco, I., Jr. (1982). Comparison of the Kinetics of Ribooligonucleotide, Deoxyribooligonucleotide, and Hybrid Oligonucleotide Double-Strand Formation by Temperature-Jump Kinetics, *Biochemistry* **21**, 5289–5295.

Nichols, R., Andrews, P. C., Zhang, P., and Bergstrom, D. E. (1994). A Universal Nucleoside for Use at Ambiguous Sites in DNA Primers, *Nature (London)* **369**, 492–493.

Nussinov, R., Tinoco, I., Jr., and Jacobson, A. B. (1982). Secondary Structure Model for the Complete Simian Virus 40 Late Precursor mRNA, *Nucleic Acids Res.* **10**, 351–363.

Okada, Y., Streisinger, G., Owen, J., Newton, J., Tsugita, A., and Inouye, M. (1972). Molecular Basis of a Mutational Hot Spot in the Lysozyme Gene of Bacteriophage T4, *Nature (London)* **236**, 338–341.

Olmsted, M. C., Anderson, C. F., and Record, M. T., Jr. (1989). Monte Carlo description of oligoelectrolyte properties of DNA oligomers: Range of the end effect and the approach of molecular and thermodynamic properties to the polyelectrolyte limits, *Proc. Natl. Acad. Sci. USA* **86**, 7766–7770.

Olson, W. K. (1982). Theoretical Studies of Nucleic Acid Conformation: Potential Energies, Chain Statistics, and Model Building in *Topics in Nucleic Acid Structure*, Part 2, Chapter 1, Neidle, S., Ed., Macmillan, London, UK.

Olsthoorn, S. M., Bostelaar, L. J., DeRooij, J. F. M., Van Boom, J. H. and Altona, C. (1981). Circular Dichroism Study of Stacking Properties of Oligodeoxyadenylates and Polydeoxyadenylate, *Eur. J. Biochem.* **115**, 309–321.

Orbons, L. P. M., van der Marel, G. A., van Boom, J. H., and Altona, C. (1986). Hairpin and Duplex Formation of the DNA Octamer d(m5C-G-m5C-G-T-G-m5C-G) in Solution. An NMR Study, *Nucleic Acids Res.* **14**, 4187–4196.

Orbons, L. P. M., van der Marel, G. A., van Boom, J. H., and Altona, C. (1987). Conformational and Model-Building Studies of the Mismatched DNA Octamer d(m5C-G-m5C-G-T-G-m5C-G). *J. Biomol. Structure. Dyn.* **4**, 965–987.

Owen, R. J., Hill, L. R., and Lapage, S. P. (1969). Determination of DNA Base Compositions from Melting Profiles in Dilute Buffers, *Biopolymers* **7**, 503–516.

Oxender, D. L., Zurawski, G., and Yanofsky, C. (1979). Attenuation in the Escherichia coli Tryptophan Operon: Role of RNA Secondary Structure Involving the Tryptophan Codon Region, *Proc. Natl. Acad. Sci. USA* **76**, 5524–5528.

Owczarzy, R., Vallone, P. M., Gallo, F. J., Paner, T. M., Lane, M. J., and Benight, A. S. (1997). Predicting Sequence-Dependent Melting Stability of Short Duplex DNA Oligomers, *Biopolymers* **44**, 217–239.

Panayotatos, N. and Wells, R. D. (1981). Cruciform Structures in Supercoiled DNA, *Nature (London)* **289**, 466–470.

Papanicolaou, C., Gouy, M., and Ninio, J. (1984). An Energy Model that Predicts the Correct Folding of Both the tRNA and the 5S RNA Molecules, *Nucleic Acids Res.* **12**, 31–44.

Pardi, A., Martin, F. H., and Tinoco, I., Jr. (1981). Comparative Study of Ribonucleotide, Deoxyribonucleotide, and Hybrid Oligonucleotide Helices by Nuclear Magnetic Resonance, *Biochemistry* **20**, 3986–3996.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K. and Breslauer, K. J. (1982). Extra Adenosine Stacks into the Self-Complementary d(CGCAGAATTCGCG) Duplex in Solution, *Biochemistry* **21**, 445–451.

Patel, D. J., Shapiro, L., and Hare, D. (1987). DNA and RNA NMR Studies of Conformations and Dynamics in Solution, *Q. Rev. Biophys.* **20**, 35–112.

Peattie, D. A., Douthwaite, S., Garrett, R. A., and Noller, H. F. (1981). A "Bulged" Double Helix in a RNA-Protein Contact Site, *Proc. Natl. Acad. Sci. USA* **78**, 7331–7335.

Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991). Internal loops in oligoribonucleotides: Symmetric loops are more stable than asymmetric loops, *Biochemistry* **30**, 6428–6436.

Petersheim, M. and Turner, D. H. (1983). Base-Stacking and Base-Pairing Contributions to Helix Stability: Thermodynamics of Double-Helix Formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp, *Biochemistry* **22**, 256–263.

Peterson, R. D., Bartel, D. P., Szostak, J. W., Horvath, S. J. and Feigon, J. (1994). ^{1}H NMR studies of the high-affinity Rev binding site of the Rev responsive element of HIV-1 mRNA: Base pairing in the core binding element, *Biochemistry* **33**, 5357–5366.

Pinnavaia, T. J., Marshall, C. L., Mettler, C. M., Fisk, C. L., Miles, H. T. and Becker, E. D. (1978). Alkali Metal Ion Specificity in the Solution Ordering of a Nucleotide, 5' Guanosine Monophosphate, *J. Am. Chem. Soc.* **100**, 3625–3627.

Pleij, C. W. A., Rietveld, K., and Bosch, L. (1985). A New Principle of RNA Folding Based on Pseudo Knotting, *Nucleic Acids Res.* **13**, 1717–1731.

Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994). Model for an RNA Tertiary Interaction from the Structure of an Intermolecular Complex between a GAAA Tetraloop and an RNA Helix, *Nature (London)* **372**, 111–113.

Podder, S. K. (1971). Co-operative Non-Enzymic Base Recognition: A Kinetic Study of Interaction Between GpGpGpC and GpCpCpC and of Self-Association of GpGpGpC, *Eur. J. Biochem.* **22**, 467–477.

Pohl, F. M. (1976). Polymorphism of a synthetic DNA in solution, *Nature (London)* **260**, 365–366.

Pohl, F. M. and Jovin, T. M. (1972). Salt-induced Co-operative Conformational Change of a Synthetic DNA: Equilibrium and Kinetic Studies with Poly(dG-dC), *J. Mol. Biol.* **67**, 375–396.

Poland, D. (1978). *Cooperative Equilibria in Physical Biochemistry*, Clarendon Press, Oxford, UK.

Pörschke, D. (1974). Thermodynamic and Kinetic Parameters of an Oligonucleotide Hairpin Helix, *Biophys. Chem.* **1**, 381–386.

Pörschke, D. (1978). Molecular States in Single-Stranded Adenylate Chains by Relaxation Analysis, *Biopolymers* **17**, 315–323.

Pörschke, D. and Eggers, F. (1972). Thermodynamics and Kinetics of Base-Stacking Interactions, *Eur. J. Biochem.* **26**, 490–498.

Pörschke, D. and Eigen, M. (1971). Cooperative Non-enzymic Base Recognition III. Kinetics of the Helix-Coil Transition of the Oligoribouridylic:Oligoriboadenylic Acid System and of Oligoriboadenylic Acid Alone at Acidic pH, *J. Mol. Biol.* **62**, 361–381.

Pörschke, D., Uhlenbeck, O. C., and Martin, F. H. (1973). Thermodynamics and Kinetics of the Helix-Coil Transition of Oligomers Containing GC Base Pairs, *Biopolymers* **12**, 1313–1335.

Powell, J. T., Richards, E. G., and Gratzer, W. B. (1972). The Nature of Stacking Equilibria in Polynucleotides, *Biopolymers* **11**, 235–250.

Prakash, G. and Kool, E. T. (1992). Structural Effects in the Recognition of DNA by Circular Oligonucleotides, *J. Am. Chem. Soc.* **114**, 3523–3527.

Praseuth, D., Perrouault, L., LeDoan, T., Chassignol, M., Thuong, N., and Helene, C. (1988). Sequence-specific binding and photocrosslinking of Z and B oligodeoxynucleotides to the major groove of DNA via triple-helix formation, *Proc. Natl. Acad. Sci. USA* **85**, 1349–1353.

Pyle, A. M. and Cech, T. R. (1991). Ribozyme Recognition of RNA by Tertiary Interactions with Specific Ribose 2'-OH Groups, *Nature (London)* **350**, 628–631.

Pyle, A. M., McSwiggen, J. A., and Cech, T. R. (1990). Direct Measurement of Oligonucleotide Substrate Binding to Wild-Type and Mutant Ribozymes from *Tetrahymena*, *Proc. Natl. Acad. Sci. USA* **87**, 8787–8191.

Pyle, A. M., Moran, S., Strobel, S. A., Chapman, T., Turner, D. H., and Cech, T. R. (1994). Replacement of the Conserved G·U with a G-C Pair at the Cleavage Site of the *Tetrahymena* Ribozyme Decreases Binding, Reactivity, and Fidelity, *Biochemistry* **33**, 13856–13863.

Pyle, A. M., Murphy, F. L. and Cech, T. R. (1992). RNA Substrate Binding Site in the Catalytic Core of the *Tetrahymena* Ribozyme, *Nature (London)* **358**, 123–128.

Quartin, R. S. and Wetmur, J. G. (1989). The Effect of Ionic Strength on the Hybridization of Oligodeoxynucleotides with Reduced Charge due to Methylphosphonate Linkages to Unmodified Oligodeoxynucleotides Containing the Complementary Sequence, *Biochemistry* **28**, 1040–1047.

Record, M. T., Jr. (1967). Electrostatic Effects on Polynucleotide Transitions. II. Behavior of Titrated Systems, *Biopolymers* **5**, 993–1008.

Record, M. T., Jr. (1975). Effects of Na^+ and Mg^{++} Ions on the Helix-Coil Transition of DNA, *Biopolymers* **14**, 2137–2158.

Record, M. T., Jr., Woodbury, C. P., and Lohman, T. M. (1976). Na^+ effects on transitions of DNA and polynucleotides of variable linear charge density, *Biopolymers* **15**, 893–915.

Record, M. T. and Zimm, B. H. (1972). Kinetics of the helix–coil transition in DNA, *Biopolymers* **11**, 1435–1454.

Rentzeperis, D., Rippe, K., Jovin, T. M., and Marky, L. A. (1992). Calorimetric Characterization of Parallel-Stranded DNA: Stability, Conformational Flexibility, and Ion Binding, *J. Am. Chem. Soc.* **114**, 5926–5928.

Richards, E. G., Flessel, C. P., and Fresco, J. R. (1963). Polynucleotides. IV. Molecular Properties and Conformation of Polyribouridylic Acid, *Biopolymers* **1**, 431–446.

Riesner, D., Maass, G., Thiebe, R., Philippsen, P., and Zachau, H. G. (1973). The Conformational Transitions in Yeast $tRNA^{Phe}$ as Studied with $tRNA^{Phe}$ Fragments, *Eur. J. Biochem.* **36**, 76–88.

Riesner, D. and Römer, R. (1973). Thermodynamics and Kinetics of Conformational Transitions in Oligonucleotides and tRNA, in *Physico-Chemical Properties of Nucleic Acids*, Vol. 2, Duchesne, J., Ed., Academic, New York, pp. 237–318.

Riley, M., Maling, B., and Chamberlin, M. J. (1966). Physical and Chemical Characterization of Two- and Three- stranded Adenine-Thymine and Adenine-Uracil Homopolymer complexes, *J. Mol. Biol.* **20**, 359–389.

Rippe, K. and Jovin, T. M. (1992). Parallel-Stranded Duplex DNA, *Methods Enzymol.* **211**, 199–220.

Rivas, E. and Eddy, S. (1999). A Dynamic Programming Algorithm for RNA Structure Prediction Including Pseudoknots, *J. Mol. Biol.* **285**, 2053–2068.

Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C., and Klug, A. (1974). Structure of Yeast Phenylalanine tRNA at 3A Resolution, *Nature (London)* **250**, 546–551.

Robinson, D. R. and Grant, M. E. (1966). The Effects of Aqueous Salt Solutions on the Activity Coefficients of Purine and Pyrimidine Bases and Their Relation to the Denaturation of Deoxyribonucleic Acid by Salts, *J. Biol. Chem.* **241**, 4030–4042.

Romaniuk, P. J., Lowary, P., Wu, H-N, Stormo, G., and Uhlenbeck, O. C. (1987). RNA Binding Site of R17 Coat Protein, *Biochemistry* **26**, 1563–1568.

Rommens, J., Kerem, B.-S., Greer, W., Chang, P., Tsui, L.-C., and Ray, P. (1990). Rapid Nonradioactive Detection of the Major Cystic Fibrosis Mutation, *Am. J. Hum. Genet.* **46**, 395–396.

Ross, P. D. and Scruggs, R. L. (1965). Heat of the Reaction Forming the Three-Stranded Poly(A+2U) Complex, *Biopolymers* **3**, 491–496.

Saenger, W. (1984). *Principles of Nucleic Acid Structure*, Chapters 9, 11, and 12, Springer-Verlag, New York.

SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991). Stabilities of Consecutive A·C, C·C, G·G, U·C, and U·U Mismatches in RNA Internal Loops: Evidence for Stable Hydrogen-Bonded U·U and $C \cdot C^+$ Pairs, *Biochemistry* **30**, 8242–8251.

SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992). Context Dependence of Hydrogen Bond Free Energy Revealed by Substitutions in an RNA Hairpin, *Science* **256**, 217–219.

SantaLucia, J., Jr. and Turner, D. H. (1993). Structure of $(rGGCGAGCC)_2$ in solution from NMR and restrained molecular dynamics, *Biochemistry* **32**, 12612–12623.

SantaLucia, J., Jr. (1998). A Unified View of Polymer, Dumbbell, and Oligonucleotide DNA Nearest-neighbor Thermodynamics, *Proc. Natl. Acad. Sci. USA* **95**, 1460–1465.

Sassanfar, M. and Szostak, J. W. (1993). An RNA Motif that Binds ATP, *Nature,(London)* **364**, 550–563.

Schera, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray, *Science* **270**, 467–470.

Schroeder, S., Kim, J., and Turner, D. H. (1996). GA and UU Mismatches Can Stabilize RNA Internal Loops of Three Nucleotides, *Biochemistry* **35**, 16105–16109.

Schmitz, M. and Steger, G. (1996). Description of RNA Folding by "Simulated Annealing," *J. Mol. Biol.* **255**, 254–266.

Sen, D. and Gilbert, W. (1988). Formation of Parallel Four-Stranded Complexes by Guanine-rich Motifs in DNA and its Implications for Meiosis, *Nature (London)* **334**, 364–366.

Sen, D. and Gilbert, W. (1990). A Sodium-Potassium Switch in the Formation of Four-Stranded G4–DNA, *Nature (London)* **344**, 410–414.

Senior, M. M., Jones, R. A., and Breslauer, K. J. (1988a). Influence of Loop Residues on the Relative Stabilities of DNA Hairpin Structures, *Proc. Natl. Acad. Sci. USA* **85**, 6242–6246.

Senior, M., Jones, R. A., and Breslauer, K. J. (1988b). Influence of Dangling Thymidine Residues on the Stability and Structure of two DNA Duplexes, *Biochemistry* **27**, 3879–3885.

Serra, M. J., Axenson, T. J., and Turner, D. H. (1994). A Model for the Stabilities of RNA Hairpins Based on a Study of the Sequence Dependence of Stability for Hairpins of Six Nucleotides, *Biochemistry* **33**, 14289–14296.

Serra, M. J., Lyttle, M. H., Axenson, T. J., Schadt, C. A., and Turner, D. H. (1993). RNA hairpin loop stability depends on closing base pair, *Nucleic Acids Res.* **21**, 3845–3849.

Serra, M. J., Barnes, T. W., Betschart, K., Guiterrez, M. J., Sprouse, K. J., Riley, C. K., Stewart, L., and Temel, R. E. (1997). Improved Parameters for the Prediction of RNA Hairpin Stability, *Biochemistry* **36**, 4844–4851.

Serra, M. J. and Turner, D. H. (1995). Predicting the Thermodynamic Properties of RNA, *Methods Enzymol.* **259**, 242–261.

Shen, P., Zengel, J. M., and Lindahl, L. (1988). Secondary Structure of the Leader Transcript from the Escherichia coli S10 Ribosomal Protein Operon, *Nucleic Acids Res.* **16**, 8905–8924.

Simpkins, H. and Richards, E. G. (1967). Preparation and Properties of Oligouridylic Acids, *J. Mol. Biol.* **29**, 349–356.

Steger, G. (1994). Thermal Denaturation of Double-stranded Nucleic Acids: Prediction of Temperatures Critical for Gradient Gel Electrophoresis and Polymerase Chain Reaction, *Nucleic Acids Res.* **22**, 2760–2768.

Steger, G., Hofmann, H., Förtsch, J., Gross, H. J., Randles, J. W., Sänger, H. L., and Riesner, D. (1984). Conformational Transitions in Viroids and Virusoids: Comparison of Results from Energy Minimization Algorithm and from Experimental Data, *J. Biomolec. Struct. and Dynamics* **2**, 543–571.

Stevens, C. L. and Felsenfeld, G. (1964). The Conversion of Two-Stranded Poly (A+U) to Three-Strand Poly (A+2U) and Poly A by Heat, *Biopolymers* **2**, 293–314.

Strobel, S. A. and Cech, T. R. (1993). Tertiary Interactions with the Internal Guide Sequence Mediate Docking of the P1 Helix into the Catalytic Core of the *Tetrahymena* Ribozyme, *Biochemistry* **32**, 13593–13604.

Strobel, S. A., Moser, H. E., and Dervan, P. B. (1988). Double-Strand Cleavage of Genomic DNA at a Single Site by Triple-Helix Formation, *J. Am. Chem. Soc.* **110**, 7927–7929.

Strobel, S. A. and Cech, T. R. (1995). Minor Groove Recognition of the Conserved G•U Pair at the *Tetrahymena* Ribozyme Reaction Site, *Science* **267**, 675–679.

Studier, F. W. (1969). Effects of the Conformation of Single-stranded DNA on Renaturation and Aggregation, *J. Mol. Biol.* **41**, 199–209.

Sturtevant, J. M. (1987). Biochemical Applications of Differential Scanning Calorimetry, *Annu. Rev. Phys. Chem.* **38**, 463–488.

Sugimoto, N., Kierzek, R., and Turner, D. H. (1987a). Sequence Dependence for the Energetics of Dangling Ends and Terminal Base Pairs in Ribonucleic Acid, *Biochemistry* **26**, 4554–4558.

Sugimoto, N., Kierzek, R. and Turner, D. H. (1987b). Sequence Dependence for the Energetics of Terminal Mismatches in Ribooligonucleotides, *Biochemistry* **26**, 4559–4562.

Sugimoto, N., Kierzek, R. and Turner, D. H. (1988). Kinetics for Reaction of a Circularized Intervening Sequence with CU, UCU, CUCU, and CUCUCU: Mechanistic Implications from the Dependence on Temperature and on Oligomer and Mg^{2+} Concentrations, *Biochemistry* **27**, 6384–6392.

Sugimoto, N., Sasaki, M., Kierzek, R., and Turner, D. H. (1989a). Binding of a Fluorescent Oligonucleotide to a Circularized Intervening Sequence from Tetrahymena thermophila, *Chemistry Lett.* 2223–2226.

Sugimoto, N., Tomka, M., Kierzek, R., Bevilacqua, P. C. and Turner, D. H. (1989b). Effects of Substrate Structure on the Kinetics of Circle Opening Reactions of the Self-splicing Intervening Sequence from *Tetrahymena thermophila*: Evidence for Substrate and Mg^{2+} Binding Interactions, *Nucleic Acids Res.* **17**, 355–371.

Sugimoto, N., Nakano, S., Katoh, M., Matsumura, A., Nakamuta, H., Ohmichi, T., Yoneyama, M., and Sasaki, M. (1995). Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes, *Biochemistry* **34**, 11211–11216.

Sullivan, K. M. and Lilley, D. M. J. (1986). A Dominant Influence of Flanking Sequences on a Local Structural Transition in DNA, *Cell* **47**, 817–827.

Sundaralingam, M. (1969). Stereochemistry of Nucleic Acids and their Constituents. IV. Allowed and Preferred Conformations of Nucleosides, Nucleoside Mono-, Di, Tri, Tetraphosphates, Nucleic Acids and Polynucleotides, *Biopolymers* **7**, 821–860.

Sundaralingam, M. (1973). Conformation of Biological Molecules and Polymers, The Concept of a Conformationally 'Rigid' Nucleotide and its Significance in Polynucleotide Conformational Analysis, *Jerusalem Symp. Quantum Chem. Biochem.* **5**, 417–455.

Sundquist, W. I. and Klug, A. (1989). Telomeric DNA Dimerizes by Formation of Guanine Tetrads Between Hairpin Loops, *Nature (London)* **342**, 825–829.

Sutcliffe, J. G. (1978), Complete Nucleotide Sequence of the *Escherichia coli* plasmid pBR322, *Cold Spring Harbor Symp. Quant. Biol.* **43**, 77

Suurkuusk, J., Alvarez, J., Freire, E., and Biltonen, R. (1977). Calorimetric Determination of the Heat Capacity Changes Associated with the Conformational Transitions of Polyriboadenylic Acid and Polyribouridylic Acid, *Biopolymers* **16**, 2641–2652.

Szewczak, A. A., Moore, P. B., Chan, Y.-L., and Wool, I. G. (1993). The conformation of the sarcin/ricin loop from 28S ribosomal RNA, *Proc. Natl. Acad. Sci. USA* **90**, 9581–9585.

Tang, C. K. and Draper, D. E. (1989). An unusual mRNA pseudoknot structure is recognized by a protein translational repressor, *Cell* **57**, 531–536.

Tibanyenda, N., DeBruin, S. H., Haasnoot, C. A. G., Van Der Marel, G. A., Van Boom, J. H., and Hilbers, C. W. (1984). The Effect of Single Base-Pair Mismatches on the Duplex Stability of d(T-A-T-T-A-A-T-A-T-C-A-A-G-T-T-G) : d(C-A-A-C-T-T-G-A-T-A-T-T-A-A-T-A), *Eur. J. Biochem.* **139**, 19-27.

Tinoco, I., Jr., Uhlenbeck, O. C., and Levine, M. D. (1971). Estimation of Secondary Structure in Ribonucleic Acids, *Nature (London)* **230**, 362–367.

Topping, R. J., Stone, M. P., Brush, C. K., and Harris, T. M. (1988). Non-Watson–Crick Structures in Oligodeoxynucleotides: Self-Association of d(TpCpGpA) Stabilized at Acidic pH, *Biochemistry* **27**, 7216–7222.

Triggs-Raine, B. L. and Gravel, R. A. (1990). Diagnostic Heteroduplexes: Simple Detection of Carriers of a 4-bp Insertion Mutation in Tay-Sachs Disease, *Am. J. Hum. Genet.* **46** , 183–184.

Trulson, M. O., Cruz, P., Puglisi, J. D., Tinoco, I., Jr., and Mathies, R. A. (1987). Raman Spectroscopic Study of Left-Handed Z-RNA, *Biochemistry* **26**, 8624-8630.

Tucker, E. E., Lane, E. H., and Christian, D. S. (1981). Vapor pressure studies of hydrophobic interactions. Formation of benzene–benzene and cyclohexane–cyclohexanol dimers in dilute aqueous solution, *J. Solution Chem.* **10**, 1–20.

Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Guild, N., Stormo, G., Gayle, M., d'Aubenton-Carafa, Y., Uhlenbeck, O. C., Tinoco, I., Jr., Brody, E. N., and Gold, L. (1988). CUUCGG Hairpins: Extraordinarily Stable RNA Secondary Structures Associated with Various Biochemical Processes, *Proc. Natl. Acad. Sci. USA* **85**, 1364–1368.

Turner, D. H. (1986). Temperature-Jump Methods, in *Investigations of Rates and Mechanisms of Reactions*, Vol. 6, Part 2, Chapter 3, Bernasconi, C. F., Ed., J. Wiley, New York.

Turner, D. H., Sugimoto, N., and Freier, S. M. (1988). RNA Structure Prediction, *Annu. Rev. Biophys. Biophys. Chem.* **17**, 167–192.

Turner, D. H., Sugimoto, N., Kierzek, R., and Dreiker, S. D. (1987). Free Energy Increments for Hydrogen Bonds In Nucleic Acid Base Pairs, *J. Am. Chem. Soc.* **109**, 3783–3785.

Uhlenbeck, O. C., Borer, P. N., Dengler, B. and Tinoco, I., Jr. (1973). Stability of RNA Hairpin Loops: A6-Cm-U6, *J. Mol. Biol.* **73**, 483–496.

van den Hoogen, Y. T., van Beuzekom, A. A., de Vroom, E., van der Marel, G. A., van Boom, J. H., and Altona, C. (1988a). Bulge-out Structures in the Single-stranded Trimer AUA and in the Duplex (CUGGUGCGG).(CCGCCCAG). A Model-Building and NMR Study, *Nucleic Acids Res.* **16**, 5013–5030.

van der Hoogen, Y. T., Erkelens, C., de Vroom, E., van der Marel, G. A., van Boom, J. H., and Altona, C. (1988b). Influence of Uracil on the Conformational Behavior of RNA Oligonucleotides in Solution, *Eur. J. Biochem.* **173**, 295–303.

Varani, G., Cheong, C., and Tinoco, I., Jr. (1991). Structure of an Unusually Stable RNA Hairpin, *Biochemistry* **30** , 3280–3289.

Voloshin, O. H., Mirkin, S. M., Lyamichev, V. I., Belotserkovskii, B. P., and Frank-Kamenetskii, M. D. (1988). Chemical Probing of Homopurine-homopyrimidine Mirror Repeats in Supercoiled DNA, *Nature (London)* **333**, 475–476.

Wada, A., Yubuki, S. and Husimi, Y. (1980). Fine Structure in the Thermal Denaturation of DNA High Temperature Resolution Spectrophotometric Studies, *Crit. Rev. Biochem.* **9**, 87–144.

Walter, A. E., Turner, D. H., Kim, J., Lyttle, M. H., Müller, P., Mathews, D. H., and Zuker, M. (1994). Coaxial Stacking of Helixes Enhances Binding of Oligoribonucleotides and Improves Predictions of RNA Folding, *Proc. Natl. Acad. Sci. USA* **91**, 9218–9222.

Walter, A. E. and Turner, D. H. (1994). Sequence Dependence of Stability for Coaxial Stacking of RNA Helixes with Watson–Crick Base Paired Interfaces, *Biochemistry* **33**, 12715–12719.

Wang, A. H.-J., Fujii, S., van Boom, J. H., and Rich, A. (1982). Right-handed and Left-handed Double-helical DNA: Structural Studies, *Cold Spring Harbor Symp. Quant. Biol.* **47**, 33–44.

Wartell, R. M. and Benight, A. S. (1985). Thermal Denaturation of DNA Molecules: A Comparison of Theory with Experiment, *Phys. Rep.* **126**, 67–107.

Weeks, K. M. and Crothers, D. M. (1993). Major Groove Accessibility of RNA, *Science* **261**, 1574–1577.

Wells, R. D., Collier, D. A., Hanvey, J. C., Shimizu, M., and Wohlrab, F. (1988). The chemistry and biology of unusual DNA structures adopted by oligopurine:oligopyrimidine sequences, *FASEB J.* **2**, 2939–2949.

Wetmur, J. G. (1991). DNA Probes: Applications of the Principles of Nucleic Acid Hybridization, *Critical Rev. Biochem. Mol. Biol.* **26** , 227–259.

Wetmur, J. G. and Davidson, N. (1968). Kinetics of Renaturation of DNA, *J. Mol. Biol.* **31**, 349–370.

Williams, A. L., Jr. and Tinoco, I., Jr. (1986). A Dynamic Programming Algorithm for Finding Alternative RNA Secondary Structures, *Nucleic Acids Res.* **14**, 299–315.

Williams, A. P., Longfellow, C. E., Freier, S. M., Kierzek, R., and Turner, D. H. (1989). Laser Temperature-Jump, Spectroscopic and Thermodynamic Study of Salt Effects on Duplex Formation by dGCATGC, *Biochemistry* **28**, 4283–4291.

Williamson, J. R. (1994). G-Quartet Structures in Telomeric DNA, *Annu. Rev. Biophys. Biomol. Struct.* **23**, 703–730.

Williamson, J. R., Raghuraman, M. K., and Cech, T. R., (1989). Monovalent Cation-Induced Structure of Telomeric DNA: The G-Quartet Model, *Cell* **59**, 871–880.

Wimberly, B., Varani, G., and Tinoco, I., Jr. (1993). The conformation of loop E of eukaryotic 5S ribosomal RNA, *Biochemistry* **32**, 1078–1087.

Woodson, S. A. and Crothers, D. M. (1987). Proton Nuclear Magnetic Resonance Studies on Bulge-Containing DNA Oligonucleotides from a Mutational Hot-Spot Sequence, *Biochemistry* **26**, 904–912.

Wu, M. J., McDowell, J. A., and Turner, D. H. (1995). A Periodic Table of Symmetric Tandem Mismatches in RNA, *Biochemistry* **34**, 3204–3211.

Wu, M., SantaLucia, J. Jr., and Turner, D. H. (1997). Solution Structure of $(rGGCAGGCC)_2$ by Two dimensional NMR and the Iterative Relaxation Matrix Approach, *Biochemistry* **36**, 4449–4460.

Wu, M. and Turner, D. H. (1996). Solution Structure of (rGCGGACGC), by Two-Dimensional NMR and the Iterative Relaxation Matrix Approach, *Biochemistry* **35**, 9677–9689.

Wyatt, J. R., Puglisi, J. D., and Tinoco, I., Jr. (1990). RNA Pseudoknots: Stability and Loop Size Requirements, *J. Mol. Biol.* **214**, 455–470.

Xia, T., McDowell, J. A., and Turner, D. H. (1997). Thermodynamics of Nonsymmetric Tandem Mismatches Adjacent to G·C Base Pairs in RNA, *Biochemistry* **36**, 12486–12497.

Xia, T., SantaLucia, J. Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998). Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson-Crick Base Pairs, *Biochemistry* **37**, 14719–14735.

Xia, T., Mathews, D. H., and Turner, D. H. (1999). Thermodynamics of RNA Secondary Structure Formation, in *Prebiotic Chemistry, Molecular Fossils, Nucleotides, and RNA*, Söll, D., Moore, P. B. and Nishimura, S., Ed., Elsevier Science Ltd., Oxford.

Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E., and Famulok, M. (1996). Structural Basis of Ligand Discrimination by Two Related RNA Aptamers Resolved by NMR Spectroscopy, *Science* **272**, 1343–1347.

Yoon, K., Hobbs, C. A., Koch, J., Sardaro, M., Kutny, R., and Weis, A. L. (1992). Elucidation of the Sequence-Specific Third-Strand Recognition of Four Watson–Crick Base Pairs in a Pyrimidine Triple Helix Motif: T·AT, C·GC, T·CG, and G·TA, *Proc. Natl. Acad. Sci. USA* **89**, 3840–3844.

Yuan, R. C., Steitz, J. A., Moore, P. B., and Crothers, D. M. (1979). The 3′ Terminus of 16S rRNA: Secondary Structure and Interaction with Ribosomal Protein S1, *Nucleic Acids Res.* **7**, 2399-2418.

Zimm, B. H. and Bragg, J. K. (1959). Theory of the Phase Transition Between Helix and Random Coil in Polypeptide Chains, *J. Chem. Phys.* **31**, 526–535.

Zimmerman, S. B., Cohen, G. H., and Davies, D. R. (1975). X-ray Fiber Diffraction and Model-building Study of Polyguanylic Acid and Polyinosinic Acid, *J. Mol. Biol.* **92**, 181–192.

Zuker, M., Jaeger, J. A., and Turner, D. H. (1991). A Comparison of Optimal and Suboptimal RNA Secondary Structures Predicted by Free Energy Minimization with Structures Determined by Phylogenetic Comparison, *Nucleic Acids Res.* **19**, 2707–2714.

Zuker, M. and Sankoff, D. (1984). RNA Secondary Structures and Their Prediction, *Bull. Math. Biol.* **46**, 591–521.

Zuker, M. and Stiegler, P. (1981). Optimal Computer Folding of Large RNA Sequences Using Thermodynamics and Auxiliary Information, *Nucleic Acids Res.* **9**, 133–148.

Zuker, M. (1989). On Finding All Suboptimal Foldings of an RNA Molecule, *Science* **244**, 48–52.

CHAPTER 9

Size and Shape of Nucleic Acids in Solution